

Cited Reference 3

특2001-0086259

(19) 대한민국특허청(KR)

(12) 공개특허공보(A)

(51) Int. Cl.⁷
G06F 17/30

(11) 공개번호 특2001-0086259
(43) 공개일자 2001년09월10일

(21) 출원번호 10-2000-7010220
(22) 출원일자 2000년09월15일
 변역문제 출원일자 2000년09월15일
(86) 국제출원번호 PCT/US1999/05588 (87) 국제공개번호 WO 1999/48028
(86) 국제출원출원일자 1999년03월16일 (87) 국제공개일자 1999년09월23일
(81) 지정국
 국내특허 : 알바니아 아르메니아 오스트리아 오스트레일리아 이제르바이
 진 보스니아-헤르체고비나 비베이도스 불가리아 브라질 벨라루스 케니
 다 스위스 리히텐슈타인 중국 쿠비 체코 독일 덴마크 에스토니아
 스페인 핀란드 영국 그루지아 헝가리 이스라엘 아이슬란드 일본 케
 니 키르기즈 북한 대한민국 키지흐스탄 세인트루시아 스리랑카 라이
 베리아 레소토 리투아니아 룩셈부르크 리트비아 몰도바 미디기스키르
 마케도니아 몽고 말리위 멕시코 노르웨이 뉴질랜드 슬로베니아 슬로
 바키아 티지키스탄 투르크메니스탄 터어키 트리니다드토비고 우크라이
 나 우간다 미국 우즈베키스탄 베트남 폴란드 포르투갈 루마니아 러
 시아 수단 스웨덴 싱가포르 아랍에미리트 안티구아바루다 코스타리카
 도미니카연방 일제리 모로코 탄자니아 남아프리카 발리즈 모잠비크
 AP-ARIPO특허 : 케냐 레소토 말리위 수단 스와질랜드 우간다 시에라리
 온 기니 감비아 짐바브웨
 EA 유라시아특허 : 아르메니아 이제르바이진 벨라루스 키르기즈 키지흐
 스탄 몰도바 러시아 티지키스탄 투르크메니스탄
 EP 유럽특허 : 오스트리아 벨기에 스위스 리히텐슈타인 독일 덴마크
 스페인 프랑스 영국 그리스 아일랜드 이탈리아 룩셈부르크 모나코
 네덜란드 포르투갈 스웨덴 핀란드 사이프러스
 OA OAPI특허 : 부르키나피소 베냉 중앙아프리카 콩고 코트디부와르 카
 메룬 기네 기네 말리 모리타니 니제르 세네갈 차드 토고 기네비소

(30) 우선권주장 60/078,199 1998년03월16일 미국(US)
09/115,802 1998년07월15일 미국(US)
(71) 출원인 엔비씨아이 뉴질랜드, 엘엘씨 알렉 인토니 이.
미합중국 캘리포니아 94104 샌프란시스코 부시 스트리트 225, 엔비씨아이 빌
딩 20층
(72) 발명자 라이안 그랜트 제임스
뉴질랜드 크리스티처치 8002 세인트 미탄 스톤로드 30에이
라이언사운들리엄
뉴질랜드 크리스티처치 8002 세튼 스트리트 7
라이언 크레이그 매튜
뉴질랜드 웰링턴 키로리알링턴 로드 110에이
문로웨이인 알리스타
뉴질랜드 크리스티처치 쿨롬보 스트리트 2/178
로빈슨 델
뉴질랜드 크리스티처치 스넬 플레이스 11
(74) 대리인 송재현, 한규환

상시청구 : 없음

(54) 인터넷 검색 엔진

요약

본 발명은 사용자의 초기 키워드 검색 기입의 결과로서 사용자에게 제공된 일반적인 웹 페이지 검색 목록으로부터 사용자가 선택한 특정 웹 페이지 목록의 결과물을 가지고 인터넷 검색 엔진 데이터베이스를 업데이트하는 방법을 제공한다. 다른 많은 사용자의 선택을 가지고 데이터베이스를 업데이트함에 따라, 데이터베이스가 소정 키워드에 관하여 가장 많이 선택되었던 웹 목록에 우선 순위를 매기도록 업데이트 될 수 있으며, 그에 따라 동일한 키워드 검색 기입을 사용한 최종 검색에서 가장 인기있는 웹 페이지 목록을 우선으로 제시한다.

대표도

도3b

명세서

기술분야

본 발명은 강화된 데이터베이스 검색이 가능한 장치 및 방법에 관한 것으로서, 특히 인터넷 검색 엔진으로서 사용되는 장치 및 방법에 관한 것이다.

배경기술

관련 정보를 입수하면서도 원하지 않거나 흥미 없는 정보는 걸러주는 효율적이고 실용적인 수단이 인터넷의 초창기부터 계속해서 요구되었다. 요즘들어 월드 와이드 웹 사이트의 수와 거기에 담겨진 정보의 양이 급증함에 따라 이러한 요구가 특히 긴박하다. 인터넷 상에서 유용한 정보를 색인하려는 시도에서, 사용자가 개인용 컴퓨터의 키보드로 적당한 키워드로 구성된 검색 명령어를 입력하게 하는 수 많은 검색 엔진 소프트웨어가 생겨났다. 이러한 검색 명령어는 서버 컴퓨터와 연결된 검색 엔진을 구비한 서버 컴퓨터에 전송된다. 검색 엔진은 검색 명령어를 받아 그것을 사용하여 웹 사이트 상에 저장된 웹 어드레스 및 텍스트의 데이터베이스를 통해 이러한 키워드를 조사한다. 그 후, 조사의 결과물이 서버 컴퓨터로부터 다시 사용자의 컴퓨터로 전송되어 사용자의 컴퓨터 스크린 상에 표시된다.

검색 엔진으로 하여금 새로운 웹 사이트를 인식하고 또한 기존의 사이트의 기록을 업데이트시키기 위해서, 그러한 웹사이트의 소유자가 이를 스스로 검색 엔진에 알리거나 혹은 서버 컴퓨터에서 데이터베이스를 업데이트시키기 위해 '웹 크롤러(web crawler)'를 통해 정보를 얻게 될 것이다. 웹 크롤러는 현재의 모든 사이트를 색인하려는 시도에서 사이트들간에 널리 흩어져서 웹 사이트 콘텐츠 및 다른 사이트와의 링크를 탐색하고 기록하는 자동화 프로그램이다.

이러한 데이터베이스 구조 및 그것의 검색 방법은 어떠한 점에서 매우 곤란하다. 인터넷이 급성장하는 결과, 웹 사이트가 자주 삭제되고, 주소가 변경되고, 업데이트되는 등의 사실에도 불구하고, 새로운 사이트의 조차도 실질적인 진류 저장분이 초래되어 검색 엔진에 이미 지난 정보 및/또는 잘못된 정보가 남아있게 된다. 비록 웹 크롤러가 그들의 위치(제목, 삽입된 링크, 주소 등)에 따라 가능성이 높은 키워드부터 우선 순위를 매기도록 구성될 수 있다 하더라도, 사용된 검색 엔진의 종류에 따라서 웹 사이트 문서의 (그 사이트 텍스트의 대부분 또는 모두를 포함하는)실제적인 부분은 여전히 조사될 것이 요구된다. 이것은 검색 엔진이 거대한 저장 용량을 필요로 하게 한다. 또한, 전형적인 키워드 검색은 과도하게 많은 양의 자료를 도출시켜 그 대부분은 사용자에게 거의 쓸모없을 것이다. 일반적으로 사용자는 그 사이트의 간단한 설명에 기초한 리스트에서 선택하고 바라는 정보를 찾을 때까지 선택된 사이트들을 탐색한다.

이러한 결과물은 검색 엔진의 표준 시일에 따라서 순위 매겨진 목록의 형태이다. 이들 표준은 검색된 텍스트 내의 전체에서 나타나는 키워드의 발생 수로부터, (전술한 바와 같이) 특정 위치에서 사용된 키워드에 가중치를 부여하는 방식까지에 걸쳐 있을 것이다. 다중의 키워드가 사용될 경우에는, 사이트는 이용가능한 여러가지 키워드의 수에 따라 순위가 매겨진다. 이러한 모든 랭킹 시스템의 기본적인 단점은 그들의 객관성(그들은 검색 엔진의 프로그램된 표준에 따라 결정된다) 및, 실제 사용자의 의견을 측정하기 보다는 특정 형태의 사이트 디자인에 대한 강조이다. 실제로 이것은 불합리한 상황을 초래할 수 있어 가장 대중적으로 사용되는 검색 엔진에 의한 비합리적인 등급 분류를 보장하려고 시도함에 있어서, 어느 디자이너가 전술한 표준에 비추어 심사숙고하여 구성한 자신의 사이트에 대해서 그 사이트의 제공, 해독력 및 콘텐츠에 순위를 입힐 수 있다.

발명의 상세한 설명

본 발명의 목적은 조작자의 두뇌력을 이용하여 전술한 종래 검색 엔진의 단점을 개선하려는 것이다.

또한 본 발명의 목적은 현존하는 소프트웨어 엔진에 비해 시스템 저장 용량 및/또는 연산에 대한 요구는 감소시키면서도 효율성, 사용 편리성 및 유효성을 강화시킨 새로운 검색 엔진을 제공하는 것이다.

또한 본 발명의 목적은 검색 데이터의 생성일 또는 갱신일을 표시함과 더불어 그 인기도 역시 함께 다양한 방식으로 표시할 수도 있도록 하는 것이다.

상술한 본 발명의 이점을 얻기 위해서, 특히 본 발명의 일 실시예는 사용자가 기입한 최초의 키워드 검색에 의한 결과를 토대로 하여 제공된 일반 웹 페이지 목록으로부터 사용자가 특정 웹 페이지 목록을 선택한 결과물로서 인터넷 검색 엔진의 데이터베이스를 업데이트하는 방법을 제공한다. 다른 많은 사용자들의 선택 결과를 토대로 하여 데이터베이스를 업데이트함으로써, 데이터베이스는 소정 키워드에 관하여 가장 많이 선택된 웹 목록에 우선 순위를 매기도록 업데이트될 수 있고, 이것에 의해 동일한 키워드를 기입한 검색 결과물에서 가장 인기있는 웹 페이지가 가장 먼저 제시될 것이다.

본 발명의 또다른 실시예에서는, 서버 컴퓨터로부터 사용자 사이트로 전송되는 목록과 함께 제공될 콘텐츠를 결정하는 방법이 제공된다. 본 실시예에서는, 다수의 다른 개발자 사이트 중의 각각으로부터 콘텐츠 목록을 얻게 된다. 각각의 콘텐츠 목록은 콘텐츠, 개발자의 신상, 및 키워드와 키워드 선택 인지를 포함한다. 그 후에, 상이한 콘텐츠 목록에 대해 공통되게 얻어진 키워드들로부터 특정한 키워드가 결정된다. 이러한 특정 키워드에 대하여, 키워드 선택 인지는 상이한 콘텐츠 목록을 사용자 사이트에 언제 전송할지를 결정하는 데 사용된다.

또 다른 실시예에서는, 사용자에게 주어진 관련 키워드의 목록으로부터 얻어진 특정 키워드에 대하여 사용자가 선택한 결과를 가지고 키워드 데이터를 업데이트하는 방법이 제공된다. 동일한 키워드와 연관된 다른 많은 사용자들의 선택 결과에 따라 데이터베이스를 업데이트함으로써, 이후에 동일한 키워드를 입력했을 때 적합한 키워드가 제일 먼저 제공되고 표시될 수 있다.

도면의 간단한 설명

본 발명의 이러한 장점 및 또다른 장점이 도면과 함께 바람직한 실시예에 대한 다음의 상세한 설명을 통해서 명확해질 것이다.

도 1은 본 발명의 전반적인 구성 중 일부를 나타내는 도면:

도 2는 검색으로의 여러가지 입력 및 서로 다른 용량의 각각에 대하여 본 발명에 따른 엔진에 의해 제공될 출력을 나타내는 도면:

도 3a 및 도 3b는 본 발명에 따라 최종 사용자에게 제공된 검색 결과물의 구성에서 웹 페이지가 선택되는 과정의 개관을 나타내는 도면:

도 4는 본 발명에 따른 상이한 웹 페이지 검색에 사용되는 데이터 세트를 나타내는 도면:

도 5는 상술된 다양한 데이터 세트 및, 다양한 입력값과 본 발명에 따라 제공되는 제시된 웹페이지 목록을 초래하는 작용을 나타내는 도면:

도 6은 본 발명에 따른 인기도 검색의 구현을 나타내는 도면:

도 7은 본 발명에 따른 꼭주 검색을 원회시키는 구현을 나타내는 도면:

도 8은 본 발명에 따른 하이-플라이어 검색(high-flyers search) 구현을 나타내는 도면:

도 9는 본 발명에 따른 무작위 검색의 구현을 나타내는 도면:

도 10은 본 발명에 따른 과거 선호 대상 검색의 구현을 나타내는 도면:

도 11은 본 발명에 따른 수집 검색의 구현을 나타내는 도면:

도 12는 본 발명에 따른 데이터 생성 검색(data created search)의 구현을 나타내는 도면:

도 13는 본 발명에 따른 맞춤 검색의 구현을 나타내는 도면:

도 14는 본 발명에 따른 그룹 동일성에 기초하는 검색 구현을 나타내는 도면:

도 15는 본 발명에 따른 키워드 제거 특성을 나타내는 도면:

도 16은 본 발명에 따라 누적 서퍼 행적 데이터를 구성하는 데 사용될 검색 결과물을 결정하는 과정을 나타내는 도면:

도 17은 본 발명에 따른 웹 페이지의 능동 제시를 나타내는 도면:

도 18은 본 발명에 따른 웹 페이지의 수동 제시를 나타내는 도면:

도 19는 본 발명에 따라 키워드를 제시하는 개관을 나타내는 도면:

도 20은 본 발명에 따라 제시된 키워드의 데이터 세트를 생성하는 방식을 나타내는 도면:

도 21은 제시된 키워드 목록이 본 발명에 따라 생성될 수 있는 여러 방식을 나타내는 도면:

도 22는 본 발명에 따른 웹 페이지 목록에 콘텐츠가 첨부되는 방법을 나타내는 도면:

도 23은 본 발명에 따른 다양한 콘텐츠 데이터 세트 및 그것을 정렬시키는 작용을 나타내는 도면:

도 24는 본 발명에 따른 다양한 콘텐츠 데이터 세트 및 그것들로부터 데이터를 선택하는 데 사용되는 작용을 나타내는 도면:

도 25는 본 발명에 따른 웹 페이지 목록 및 다른 콘텐츠 데이터를 나타내는 도면:

실시예

도 1a 및 도 1b는 본 발명의 전반적인 구성의 일부를 나타내며, 이후에 더욱 상세히 서술될 것이다. 여러 도면과 설명에서 유사하게 매겨진 참조 번호는 하기의 설명에서 동일하거나 유사한 구조, 작용 및 과정 단계를 언급하도록 사용될 것이다.

본 발명은 바람직하게는 잘 알려진 비외 같이, 각각의 컴퓨터가 일반적으로 마이크로프로세서, 메모리 및 모뎀을 포함하며, 어떤 컴퓨터는 디스플레이 등도 포함하는 컴퓨터들의 네트워크 환경에서 구현된다. 도 1b에 도시된 비외 같이, 복수의 사용자 사이트/컴퓨터(100A-100D)가 복수의 서버 컴퓨터(102A, 102B) 및 개발자 사이트/컴퓨터(104A, 104B)와 함께 도시된다. 일반적인 인터넷 네트워크에서, 도시된 비외 같이 상이한 서버 컴퓨터(102)가 함께 상호연결될 수 있다. 또한, 비록 사용자 사이트, 개발자 사이트 및 서버 컴

퓨터가 몇 개만 도시되었지만, 그러한 컴퓨터가 무수히 상호연결되어 있다.

비록 본 발명의 특정 실시예가 도시된 바와 같이, 본 발명이 서버 컴퓨터에 의해 작동되는 코드화된 프로그램 명령의 시퀀스로서 구현되는 응용에 대하여 기술되지만, 이러한 프로그램 명령의 특정 시퀀스는, 이후에 서술되는 응용에 대해 특정하게 맞추어진 특정 명령을 구비하는 프로세서와 같은, 다른 형태에 의해 구현될 수 있음이 명확하다.

후술되는 바와 같이, 사용자에게는 보여지지 않지만 현재 행해진 검색을 바탕으로 미래의 검색 결과물을 얻기 위한 부가적인 동작들이 구현된다. 도시된 바와 같이, 본 발명은 다양한 기능을 가지며, 각각의 기능은 도시된 도 1A에서 나란히 위치하여 도시되며, 상기 도면은 상이한 기능들이 동시에 진행될 수 있음을 개관하여 도시하고 있다. 전체적인 기능에 관하여, 시작 블록(10)은 세 가지 즉, 웹 페이지 제시(12), 키워드 제시(14) 및 콘텐츠 제시(16)로 나타난다.

본 발명에 따라 사용자에게 의해 웹 페이지(12)가 선택되기 위해서는, 실행될 검색의 종류가 선택되는 단계(18)가 있다. 그 후에는, 단계(20)에서 다양한 소스 중의 하나로부터 얻어진 검색 입력값이 입력되어, 단계(18)에서 검색 결과물을 선택하기 위해 선택된 알고리즘과 함께 사용된다. 다음에는 이러한 검색의 결과물이, 도 1에서 각각의 단계(22, 24, 26)로서 도시된 바와 같이, 생성된 웹 페이지 목록을 표시하는 단계와, 제시된 웹 페이지를 수동적으로 표시하는 단계와, 제시된 웹 페이지를 능동적으로 표시하는 단계로 사용자에게 표시된다. 이러한 기능 및 그것을 이행하는 방법이 이후에 더욱 상세히 서술될 것이다.

키워드 제시를 발생(사용자는 선택할 수도 안할 수도 있다)시키기 위해서, 사용할 키워드 검색 알고리즘의 종류가 선택되는 초기 단계(28)가 있는 것이 바람직하다. 비록 많은 시스템에서는 그러한 알고리즘을 단지 하나만 구비할 터이지만, 후술되는 바와 같이 다양한 알고리즘도 가능하다. 일단 키워드 검색 알고리즘이 선택되면, 후속 단계(30)에서는 사용자가 입력한 키워드를 기초로 하여, 관련 키워드를 결정하기 위해 현재의 키워드 데이터의 세트가 연산된다. 이러한 연산의 결과물은 그 후 단계(32)에서 사용자에게 표시된다. 이러한 기능 및 그것을 이행하는 방법이 이후에 더욱 상세히 서술될 것이다.

전술한 웹 페이지 및 키워드 선택 기능은 최종 사용자에게 직접적으로 도움을 준다. 최종 사용자에게 간접적으로 효과를 주는 본 발명의 또다른 신규한 특성은 그것이 다양한 표준을 기초로 하여 실시간에 목표하는 콘텐츠에 이르게 하기 때문에, 광고주의 이익과 직결된다. 이후에 좀 더 자세하게 서술되는 바와 같이, 먼저 단계(34)에서, 콘텐츠를 선택하는 방법의 결정하기 위한 콘텐츠 제공 알고리즘이 선택된다. 후속 단계(36)에서는 사용자와 콘텐츠 제공자로부터의 입력값에 기초하여 어떠한 콘텐츠를 보여줄 것인지가 결정된다. 그런 다음, 키워드 및/또는 웹 페이지가 표시됨과 동시에, 사용자가 볼 수 있도록 광고가 표시된다.

비록 도 1이 본 발명에 따른 특정한 전반적인 특성을 도시하고 있지만, 본 발명의 많은 이로운 특성은 전술한 바와 같이 사용자에게는 보여지지 않는다. 하지만 그러한 것들은 본 발명이 구현되는 방법을 완전히 설명하기 위해 중요하고, 이하에 서술된다.

도 2는 본 발명에 따른 검색 엔진에 입력되는 다양한 입력값과, 상이한 기능에 대하여 제공될 출력값을 나타낸다. 이후에 더욱 상세한 설명이 주어진다. 검색 엔진 사용자로써 잠재적으로 입력되는 데이터는 다음과 같은 것을 포함한다.

- 키워드(52)-웹 페이지 목록을 찾기위해 사용자가 기입하는 단어 또는 어구
- 신상정보 타입(54)-예를 들어, 미국, 남성, 의사 등, 그 사람의 신원과 결부된 사람들의 집단
- 사용자 ID(56)-검색 엔진에 등록할 것을 선택한 사용자 특유의 신원 식별(identification)이며, 이는 쿠키(cookie) 또는 로그온(logon)을 통해 수행될 수 있음
- 검색 타입(58)-검색자에 의해 자신의 기호(인기도, 신규성 등)에 따라 검색 결과물의 타입을 결정하기 위해 능동적으로 선택될 수 있음
- 날짜와 시간(60)-검색자가 시스템을 사용할 때 수동으로 기록됨
- IP 어드레스(62)-검색자가 시스템을 사용할 때 수동으로 기록됨
- 기티(64)-검색 습관화 선호도, 웹 페이지 제시용 키워드 등의 다른 개인 정보를 포함하며, 이 정보가 일단 사용자 스스로에 의해 능동적으로 기입되기만 하면, (사용자 ID로 인증된) 사용자가 검색 엔진을 사용할 때마다 검색 결과물을 개인화하는데 이용됨.

웹 페이지 개발자로부터의 데이터는 다음과 같은 것을 포함한다.

- URL(66)-개발자가 제출하기를 바라는 웹 페이지의 URL 어드레스
- 설명(68)-개발자의 웹 페이지에 관한 정보에 대한 2 내지 3 줄의 설명
- 키워드(70)-웹 페이지 개발자가 자신의 웹 페이지와 연관시키고자 하는 키워드
- 목표 대상자(72)-웹 페이지 개발자가 특히 목표로 하고자 하는 목표 대상자(신상정보 타입(54))
- 날짜와 시간(74)-웹 페이지 개발자가 웹 페이지를 제출할 때마다 수동으로 기록됨.

콘텐츠 제공자로부터의 데이터는 다음과 같은 것을 포함한다.

- 입찰가(bid)(76)-후술되는 바와 같은 콘텐츠에 대한 책정가

- 세부 콘텐츠(70)-어드레스, 세부 콘텐츠 등을 포함한 콘텐츠 제공자의 모든 세부 사항을 포함
- 다른 검색 엔진으로부터의 결과물(80)-다른 데에 있는 검색 엔진으로부터의 키워드 검색에 대한 결과물.

검색 엔진(10)의 출력값은:

- 웹 페이지의 목록(90)-입력 데이터에 따라, 후술되는 웹 페이지 결정 단계(82)에서 웹 페이지 목록이 생성될 수 있음;

- 콘텐츠 키워드(92)-검색 엔진은 사용지기 후술되는 키워드 결정 단계(84)에서 생성된 다른 키워드를 시도할 것을 제안함;및

- 콘텐츠(94)-검색 엔진이 후술되는 콘텐츠 결정 단계(86)에서 생성된 선택된 콘텐츠를 전송함;

이다.

참고를 용이하게 하고 이해를 돕기 위해서, 본 발명의 비limit적인 실시예와 관련하여 상기 및 하기에서 언급된 데이터 세트의 정의가 아래에 같이 반복되고 확장된다(그리고 적절한 곳에서는, 정의된 데이터 세트를 생성하는 데 사용되는 종속된 데이터 세트가 테이블의 형태로 보여진다). 이후에, 소정의 데이터 세트 정의는 본 발명을 행하고 사용하는 방법을 전체적으로 알리기 위해 더 세부적으로 상세히 설명될 것이다.

위치: 복수의 유일한 정보 엔터티

웹 페이지: 웹 페이지 URL(Universal Reference Locator) 어드레스 형식의 위치

키워드: 검색 엔진에 기입되는 단어 또는 어구

히트 목록: 키워드 검색의 결과물인 웹 페이지(URL 어드레스)의 목록. 이러한 히트 목록은 키워드와 관련된 웹 페이지의 관련성으로 순위가 매겨진다. 히트 목록은 항상 그것과 연관된 키워드를 구비한다.

입력 데이터 세트	출력 데이터 세트
<ul style="list-style-type: none"> • 키워드 (일시적) • 키워드와 꼭 맞는 데이터베이스 (영구적) 	<ul style="list-style-type: none"> • 히트 목록 - 웹 페이지의 순위가 매겨진 히트 목록 (일시적)

영구적 데이터 세트: (시간이 경과하여 그것이 바뀌더라도)오랜 기간 유지됨.

일시적 데이터 세트: 오로지 검색하는 동안에만 생성됨.

서퍼 행적(surfer trace): 이것이 사용지의 검색식을 측정하는 지표가 된다. 이것은 사용지들이 검색하는 키워드의 행적으로서, 결과적으로 어떠한 URL이 선택되고 거기 거기서 얼마나 머물렀는가를 통해 사용지(서퍼)를 위한 웹 페이지의 순위가 계산될 수 있다. 이것은 키워드 검색 후에 사용지들이 어떤 웹 페이지를 가장 유용하다고 여기는가에 대한 웹 페이지에 대한 측정이다. 모든 서퍼 행적의 조합은 사용지의 선택 히트 목록을 생성하는 데 사용된다.

입력 데이터 세트	출력 데이터 세트
<ul style="list-style-type: none"> • 키워드 (일시적) • 초기 검색결과로 부터의 사용자 선택(일시적), 즉 방문한 웹 페이지(URL) • 각 URL 소모시간 • 사용자의 IP 어드레스 	<ul style="list-style-type: none"> • 서퍼 행적 - 각 키워드에 대해 사용자들이 유용하다고 여기는 사용자 웹 페이지 목록 (영구적일 수도 있고 일시적일 수도 있음)

사용지의 선택 히트 목록: 모든 키워드와 연관된 웹 페이지의 반영구적 순위이며, 그 키워드와 연관된 각각의 웹 페이지가 인터넷 사용자에게 얼마나 유용하게 여겨지는지에 대한 지침. 상기 사용지의 선택 히트 목록은 새로운 서퍼 행적에 의해 중립적으로 업데이트된다.

입력 데이터 세트	출력 데이터 세트
<ul style="list-style-type: none"> 서퍼 행적 - (영구적일 수도 있고 일시적일 수도 있음) 사용자의 선택 히트 목록(영구적)* 	<ul style="list-style-type: none"> 신규 사용자의 선택 히트 목록 - "인기" 웹 페이지의 순위 히트 목록 (영구적)

초기 사용자의 히트 목록은 서퍼 행적이 될 것이다.

신규 웹 페이지 목록: 웹 페이지 개발자로부터의 URL 제출에 의해 생성된 신규 웹 페이지의 목록. 웹 개발자가 웹 페이지를 업데이트할 때, 그들은 웹 페이지 어드레스, 상기 페이지에 대한 간략한 정보 및 개발자의 결정과 관련된 키워드의 목록을 제출할 수 있다.

입력 데이터 세트	출력 데이터 세트
<ul style="list-style-type: none"> 웹 어드레스 및 키워드에 대한 모든 웹페이지 개발자 정보 	<ul style="list-style-type: none"> 신규 웹 페이지 목록(영구적)

컨텐츠 제공자의 목록: 키워드를 가진 컨텐츠를 표시하는 데 일반적으로 지불해야 하는 컨텐츠 제공자의 (각각의 키워드와 관련된) 목록. 지불되는 가격은 다른 컨텐츠 제공자의 수, 개발자가 들인 비용 및 키워드가 검색된 회수에 따른다.

입력 데이터 세트	출력 데이터 세트
<ul style="list-style-type: none"> 키워드 컨텐츠 스코트에 대한 내용 제공자의 책정가 	<ul style="list-style-type: none"> 컨텐츠 제공자 목록 - 각 키워드와 관련된 컨텐츠의 목록(영구적)

하이-플라이어(high-flyers) 히트 목록: 최고 속도로 인기도가 상승하는 (모든 키워드와 관련된) 웹 페이지의 목록. 이것은 웹 페이지가 사용자의 선택 히트 리스트에서 얼마나 급속도로 상승하는지에 대한 지표이고, 새롭게 출현하는 웹 페이지가 사용자의 선택 히트 리스트의 순위로 부상하는 것을 보장하는 수단으로 이용된다.

입력 데이터 세트	출력 데이터 세트
<ul style="list-style-type: none"> 기존 사용자의 선택 히트목록 - (일시적) 신규 사용자의 선택 히트목록 - (영구적) 	<ul style="list-style-type: none"> 하이-플라이어 히트목록: 가장 빠르게 인기가 오른 웹페이지의 순위목록

개인별 히트 목록: 사용자 개개인이 과거에 행하였던 각각의 키워드 검색에서 가장 유용하다고 여겨진 웹 페이지의 목록. 이것은 사용자 개개인에게 있어 지동 북미킹(book-marking) 데이터 세트와 같은 것이다.

입력 데이터 세트	출력 데이터 세트
<ul style="list-style-type: none"> 키워드 개별적인 서퍼 행적 - (영구적) 	<ul style="list-style-type: none"> 개인별 히트목록: 개개인이 과거에 유용하다고 여긴 웹 페이지의 순위목록

수집 검색 히트 목록: 상기 히트 목록 중의 임의의 것들의 조합일 수 있음. 그러한 히트 목록이 조합되는 데에는 여러가지 상이한 방법들이 있을 수 있다.

입력 데이터 세트	출력 데이터 세트
<ul style="list-style-type: none"> • 크롤러 히트목록(일시적) • 사용자의 선택 히트목록(영구적) • 광고주의 목록(영구적) • 신규 웹페이지 목록(영구적) • 하이-플라이어 목록(영구적) • 개인별 히트목록(영구적) 	<ul style="list-style-type: none"> • 수집 검색 히트목록 - (디폴트) 키워드 검색후에 사용자에게 표시된 웹페이지의 순위 목록. 이것은 상기 히트목록 중의 임의의 것들의 조합일 수 있다.

크롤러 키워드 리스트: 사용자의 유용하게 여길 키워드 제시 목록. 이것은 다른 사용자가 시도하였던 키워드 및 어구의 데이터베이스에 사용자의 의해 가입된 키워드를 매칭시킴으로써 찾아진다. 비록 이것이 웹 페이지의 순위이기 보다는 키워드의 순위이지만, 이것은 크롤러 히트 목록과 동등하다. 이것은 철자 검사기(spell-checker)와 유사한 알고리즘을 사용하여 어구에 대하여인 철자 검사를 행하는 방법이다. 또한 이것은 사용자 키워드의 시퀀스로부터의 중전 URL 선택에 기초하여 키워드를 제시한다.

입력 데이터 세트	출력 데이터 세트
<ul style="list-style-type: none"> • 키워드(일시적) • 사용된 모든 키워드의 데이터 베이스 (영구적) 	<ul style="list-style-type: none"> • 사용자가 시도하고자 하는 다른 키워드의 순위목록(일시적)

서퍼 키워드 목록: 사용자 개개인이 키워드가 선택된 후에 유용하다고 여기는 키워드의 목록으로 구성된 데이터 세트. 이것은 사용자가 어떤 키워드를 사용하기로 결정했는지를 추적함으로써 밝혀진다. 이것은 서퍼 행적과 동등하다.

입력 데이터 세트	출력 데이터 세트
<ul style="list-style-type: none"> • 키워드(일시적) • 키워드 제시자로부터 사용되었던 키워드가 무엇인지에 대한 데이터 	<ul style="list-style-type: none"> • 개별 사용자가 유용하게 여겼던 (키워드와 관련된) 다른 키워드 순위목록(반영구적)

키워드 제시자: 사용자가 유용하다고 여겨서 연속적인 서퍼 키워드 목록으로부터 축적하고 각각의 키워드를 목록으로 만든, 다른 키워드의 영구적인 순위를 포함하는 데이터 세트(이것은 사용자의 선택 히트 목록과 동등하다).

입력 데이터 세트	출력 데이터 세트
<ul style="list-style-type: none"> • 서퍼 키워드 목록 (일시 또는 영구적) • 기존 사용자의 선택 히트 목록 (영구적) 	<ul style="list-style-type: none"> • 신규 사용자의 선택 키워드 목록(영구적)

사용자 기초 검색 알고리즘을 기초로 함.

상기 제공된 논의는 본 발명을 보다 완전하게 서술하는 데 필요한 언어를 제공한다. 도 3a 및 도 3b에 도식된 것은, 최종 사용자에게 제공된 검색 결과물을 구성하는 데 있어서 웹 페이지가 선택되는, 본 발명에 따른 검색 엔진 기능에 대한 개관을 제공한다. 단계(112)에서, 사용자는 4개 까지의 데이터 세트(키워드(52), 신상정보 타입(54), 검색 타입(58) 및 사용자 ID(56))를 가입한다. IP 어드레스(62) 및 날짜와 시간(60)은 사용자의 의해 입력되지 않고 사용자가 검색 엔진을 사용할 때 얻어질 수 있다. 이러한 데이터는 단계(114) 및 단계(116)에서 웹 페이지의 목록을 작성하는 데 동시에 사용된다. 이후에 상세히 논의되는 단계(114)는 본 발명에 따라 생성되는 신규한 새로운 검색 엔진 데이터 세트로부터 웹 페이지를 선택하는 과정이다. 필요하다면, 이 단계를 다른 현존하는 검색 엔진으로부터 웹 페이지의 선택을 얻는 단계(116)와 병행하여 실행할 수 있다. 그 후에, 단계(114) 및 단계(116)로부터의 웹 페이지 선택은 단계(118)에서 결합되고 태그(tag)된다. 하기에 더 상세히 서술되는, 웹 페이지의 목록을 태그하는 과정은 도 3에서 서퍼

행적 데이터로 도시된 데이터의 일 세트가 생성될 수 있게 하며, 또한 검색 엔진 사용자기 단계(120)에서 목록으로부터 웹 페이지를 선택할 때 이를 검색 엔진으로 되돌려 보낼 수 있도록 한다. 검색 엔진 데이터 세트를 업데이트하는 데 사용되는, 태그된 웹 페이지를 선택하는 과정은 다음의 일련의 데이터를 생성한다: 키워드(124), URL(126), 사용지 ID(128), IP 어드레스(130), 날짜와 시간(132), 관련한 웹 페이지 설명(134).

비록 서버 행적 데이터에서 이러한 모든 상이한 데이터 타입을 사용하는 것이 바람직하지만, 이 데이터들을 상이하게 조합하여 사용하는 것 역시 전체적으로 본 발명의 의도된 범위 내에 있다. 상기 설명(134)은 일반적으로 신규 사이트가 검색 엔진(10)의 데이터 세트(114)에 첨가될 때에만 본 발명의 바람직한 실시예에 포함되고, 사용된 설명은 웹 페이지의 최초 목록 상에 나타나는 설명이 될 것이다. 날짜와 시간(132)은 사용자기 특정 사이트에 머무른 시간을 기록하는 것이 아니라, 차후에 설명되는 비외 같이 사이트가 선택되었던 사실만을 표시할 수도 있다. 이러한 과정은 사용자에게는 보이지 않아, 웹 페이지의 목록으로부터 웹페이지를 선택하면 단계(122)에서 해당 URL로 직접 옮겨진다. 단계(114), 단계(118) 및 단계(120)에 대한 구체적인 구현이 이후에 더욱 상세히 서술될 것이다.

초기 선택 후 사용자는 또다른 웹 페이지 URL 검색 결과물을 액세스할 것을 접근하려 할 수 있다. 사이트의 관련성에 따라, 사용자는 원기, 내려받기, 다른 페이지, 걸려있는 링크의 탐색 등으로 시간을 소모할 것이며, 또는 사이트가 무관하거나 관심없는 것이라면 사용자는 잠깐 머물렀다가 곧바로 다시 검색 결과물로 되돌아갈 것이다. 두 선택간의 시간차는 웹 페이지 검색 목록으로부터의 순차적인 선택에 의한 두 개의 날짜/시간 데이터(132)간의 차이로서 기록된다(본 실시예에서, 또 다른 선택이 한 웹 페이지를 방문한 후에 한 것이라면 그 웹 페이지에 머무른 시간만을 측정할 수 있다-이것은 시간차가 계산되도록 할 수 있는 또다른 서버 행적(132)을 제공한다). 웹 페이지의 인기도 상의 이러한 서버 행적 데이터는 차후에 서술되는 비외 같이 후속의 검색을 순위 매김하는 데 사용된다.

따라서, 본 발명에 따르면, 검색의 주제(대상)에 대한 상이한 결과물의 관련성을 확립하는 데 사용되고 있는 것은 인간 사용자의 이성력과 분석력이다. 본 발명은 누적 과정과 모든 인간 사용자의 이성을 이용하여, 상기 서술된 현재 가능한 방법의 형태를 가진 것보다 더욱 효과적으로 원하는 정보 소스를 얻는 방법을 제공한다.

상술한 비외 같이, 인간의 두뇌력은 각 키워드 검색 후에 사용자기 방문하는 웹 페이지에 대한 기록에 의하여 포획된다. 본 발명에 따르면, 서버 행적 데이터에 대한 수집은 검색에 의해 사용자에게 전송된 웹 페이지의 목록에서 검색 엔진(또는 보조 서버)에 되돌아온 정보를 자동적으로 보낼 숨은 링크(hidden link)에 의해 수행된다. 사용자가 오로지 그의 의도대로 표시된 링크만을 보는 동안, 상기 숨은 링크는 검색 엔진에 전송을 알리며 이는 자바 애플릿(Java applet)으로 실행될 수 있는 과정이다. 따라서, 인터넷 사용자기 웹 페이지를 선택하면 사용자기 그 어드레스로 길 뿐만 아니라, 무엇이 선택되었는지를 알리는 서버 행적 데이터가 검색 엔진(10)으로 전송된다. 사용자가 웹 페이지 목록으로 돌아와 또다른 웹 페이지 목록을 선택할 때에는, 또 다른 서버 행적을 생성하도록 또 다른 자바 애플릿이 실행된다. 두 개의 순차적인 선택으로부터 이러한 서버 행적 내에서의 데이터 시간 데이터 사이의 차이는 사용자기 종전 웹 사이트에 머물렀던 시간을 획득한다. 이것은 사용자기 이러한 데이터가 전송되었는지 알지 못한 상태로 발생한다.

또다른 실시예에서는, 서버 행적 데이터의 완성된 목록을 수집하기 위해서 다중 자바 애플릿을 사용하는 것이 아니라, 설명 데이터(134)와, 사용자기 특정 웹 사이트를 방문했다는 것을 나타내는 날짜와 시간 데이터(132)가 존재하지 않는다. 특정의 일 실시예에 있어서는, 일 본 또는 오 본과 같이 소정의 시간 이상의 기간 동안 사용자기 특정 웹 사이트를 방문해야 하며, 그 시간은 이후에 서술되는 비외 같이 사이트를 방문하여 그 사이트를 관찰하고 또한 임의의 서버 행적 데이터가 검색 엔진(10)으로 되돌려 보내지는 데 적절한 시간이 얼마인가에 따라 결정된다. 본 실시예에서는, 각 애플릿이 검색 엔진의 데이터베이스를 업데이트하는 데 필요한 모든 정보를 포함한다. 또다른 실시예는 사용자기 의도한 웹 사이트를 방문하기 이전에 서버 행적 데이터를 수집한다. 본 발명이 의도하는 범위 내에서 이러한 서버 행적 데이터를 획득하는 또다른 방법이 있을 수 있다.

따라서, 본 발명에 따른 검색 결과물 페이지는 종래의 검색 엔진의 결과물 페이지와 상이한 형태(format)가 된다. 그 차이는 컨텐츠보다는 작용(action)에 있다. 시각적으로는, 사용자에게 있어서 상기 페이지가 다른 검색 엔진으로부터의 표준 검색 결과물과 같이 동일하게 보여진다.

일례로 다음과 같은 점을 들 수 있다. 종래의 검색에서, 키워드 '날씨'의 검색 결과물 페이지는 다음과 같을 것이다: 1.www.weather.com 오늘의 날씨 예보. 오늘은 전국이 맑고 화창할 것으로 기대됩니다.

'www.weather.com'의 표식과 연관된 HTTP 링크는 'http://www.weather.com'이다. 이것은 만일 사용자기 상기 링크를 선택하면, 사용자는 곧바로 이 페이지에 방문할 것이라는 것을 의미한다.

이외 반대로, 본 발명에 따르면, 키워드 '날씨'의 요구에 대한 검색의 첨부된 결과물 페이지는 다음과 같이 보일 것이다.

1.www.weather.com 오늘의 날씨 예보. 오늘은 전국이 맑고 화창할 것으로 기대됩니다.

상기 'www.weather.com'의 표식과 연관된 HTTP 링크는 'link.asp?n=1.'이다. 따라서, 만일 사용자기 상기 링크를 선택하면, 사용자에게는 보이지 않는 과정이지만, 사용자는 제일 먼저 본 발명에 따른 검색 엔진(10)을 사용하는 서버 컴퓨터에 대응하는 사이트 상의 link.asp 페이지를 먼저 항하게 되며, 파라미터 n에 값 1을 전달한다.

서버측 코드(웹 서버 상에서 실행되는 응용 코드)는 상기 파라미터를 이용하여 사용자기 고른 사이트의 URL 및 설명을 식별한다. 그 후 이러한 정보는 다른 서버 행적 데이터와 함께 데이터베이스 테이블에 저장된다. 다음에 서버측 코드는 사용자의 요구된 URL로 재전송하는 작동을 수행한다. 그러면 사용자는 원하는 페이지를 볼 수 있다.

검색 결과물의 소스는 이러한 작용과는 무관하다. 사용자의 목적 페이지는 이러한 작용과 무관하다. 이러한 과정은 사용자, 키워드 및 목적지를 데이터베이스 내에 기록하는 것 중의 하나이다. 이러한 추적 방법

은 키워드 검색 후에 방문된 최초 웹 페이지만을 기록할 수 있다. 만일 사용자기 검색 결과물 리스트로 되돌아 오기를 계속한다면 후속의 웹 페이지 방문이 기록될 수 있다.

사용자기 웹 페이지를 검색한 결과로서 검색 엔진(10)의 데이터 세트(114)로 되돌아온 서버 행적 데이터는 부정된 사용자기 검색 엔진에 허위 데이터를 전송하는 것을 막기 위해 암호화될 수 있다.

사용자기 최초 URL 선택으로부터 어디로 연결되는지를 추적하는 또 다른 방법은(사용자기 검색 결과물 페이지로 돌아오지 않는다는 가정 하에) 선택된 웹 페이지를 검색 엔진 웹 사이트에 위치한 '프레임'의 부분으로서 실행하는 것이다. 이것은 키워드가 입력된 후에 방문된 웹 페이지의 완전한 기록이 기록될 수 있게 한다. 하지만 이것은 시스템에서의 응답 시간을 감소시킬 수 있는 가능성을 갖지만, 시스템에 복잡성의 정도를 증대시킨다.

전술한 비외 같이, 수집될 수 있는 서버 행적 데이터는 키워드(124), URL(126), 사용자 ID(128), IP 어드레스(130), 날짜와 시간(132) 및 간략한 웹 페이지 설명(134)을 포함하며, 이는 검색자(서퍼)가 검색 엔진을 어떻게 사용하는지에 대한 행적 또는 기록을 제공하기 때문에 서버 행적 데이터라 불린다. 이러한 데이터는 중전 검색자의 선호도를 기초로하여 장래의 검색을 개선하는 데 이용된다. 따라서 서버 행적은 특정한 일련의 키워드에 대한 초기의 검색 결과물로부터 개별 사용자 또는 웹 서버의 더 나은 선택을 측정하는 수단이다.

웹 페이지의 목록을 결정하는 데이터 세트가 생성되는 방법.

도 4는 본 발명에 따라 상이한 웹 페이지 검색을 위해 사용된 데이터 세트를 나타낸다. 데이터 세트(테이블)는 웹 페이지가 키워드 테이블(164), 신상정보 ID 테이블(166), 보안 테이블(168), 누적 서버 행적 테이블(170), 키워드 URL 링크 테이블(172), 개인별 링크 테이블(174) 및 웹 페이지(URL) 테이블(188)을 포함하는 리스트인지 여부를 결정하는 데 사용된다.

상술한 데이터 세트의 구조는 이후에 더욱 상세히 서술된다. 이어지는 다음의 설명에서는 데이터가 스프레드시트(spreadsheet) 식으로 배열되고, 각 셀마다 다중치를 가지며 많은 공백 값을 가진다. 이러한 방식의 묘시는 본 발명을 설명하기에는 편리하지만, 저장하고 검색하는 방법으로는 효율적이지 않다. 당업자에게는 명확해질 터이지만, 본 발명에 따른 데이터 저장 수단의 구현을 위해서는 관계형 데이터베이스 모델이 사용될 것이며, 데이터 저장에 다중 필드 또는 테이블이 관련될 것이며 각 필드는 하나의 값을 저장할 것이다.

키워드 테이블(164)

도 4의 키워드 데이터 테이블(164)의 콘텐츠는 어구와 요청된 회수를 포함한 키워드의 목록으로서 아래에 도시된 테이블 1에서 좀더 상세하게 나타낸다. 목록이 더러기 힘들 정도로 커지면, 소정의 기간 후에 재사용되지 않는 키워드는 목록에서 삭제될 수 있다. 하지만 가능한 한 입력된 대다수 또는 모든 키워드 어구를 유지하는 것이 바람직하다.

키워드	키워드가 요청된 누적 회수(W)	각 키워드에 대한 고유번호
키워드 1	W1, W2, W3등	
키워드 2		
키워드 3		
키워드 4		
키워드 5		
키워드 6		
키워드 7		

테이블 1 정보요청 목록 및 그것을 요청한 회수

키워드가 요청된 누적 회수는 여러가지 선택된 '사용자 신상정보'(W1, W2, W3, ...). 예를 들어, W1 = 총 검색, W2 = 남성 신상정보, W3 = 여성 신상정보, W4 = USA 신상정보 등에 따라 분리해도 된다. 사용자는 하나 이상의 신상정보 범주, 예를 들어 USA(W3) 출신의 남성(W2)에 들어 있을 수도 있기 때문에, W의 합계는 사이트를 방문해온 사이트의 총 회수보다 클 것이라는 점에 주의해야 한다. 이것은 키워드를 사용하는 검색자 수의 목록 뿐만 아니라, 그 키워드에 대해 검색하는 (선택된 상기 신상정보 타입에 따른) 사용자 타입의 목록이 된다. 서로 다른 언어로 된 동일한 것을 의미하는 키워드는 비록 이하에 기술될 키워드 제시자를 이용하여 관련될 수 있을지라도, 철자가 다른 한, 상이한 키워드이다.

웹 페이지 테이블(188)

인터넷 웹 페이지의 목록을 포함하는 도 4의 웹 페이지 테이블(188)의 콘텐츠는 이하에 도시된 테이블 2에 좀더 상세하게 나타낸다. 각 웹 페이지는 URL 어드레스, 관련된 2-3 라인 설명, 각 URL(임의의 문자, 심볼 코드 또는 표시도 가능함)에 관한 고유 웹 페이지 번호 및 상기 URL을 방문해온 누적 회수를 갖는다. 상기

URL 어드레스는 아래의 데이터 테이블에 모든 URL 스트링을 저장하기 보다는 그것에 할당된 고유번호(임의의 문자, 심볼 코드 또는 표시도 기능)를 갖게 될 것이다.

어드레스	2-3줄의 설명	각 URL 어드레스에 대한 고유번호	URL(웹 페이지)을 방문한 빈도
URL 어드레스 1			
URL 어드레스 2			
URL 어드레스 3			
URL 어드레스 4			
URL 어드레스 5			
URL 어드레스 6			
URL 어드레스 7			

테이블 2 정보제공자의 목록 및 웹 페이지의 설명

키워드 URL 링크 테이블(172)

도 4의 키워드 URL 링크 테이블(172)의 콘텐츠는 이히에 도시한 테이블 3에서 좀더 자세하게 나타낸다. 이 테이블은 정보 제공(URL 어드레스 또는 웹 페이지)과 정보 요청(키워드)간의 링크에 대한 정보를 포함하고 있기 때문에 본 발명에 관해서 특히 중요하다.

이 데이터는 다음의 세 개의 파라미터에 의해 정의된 키워드 및 시간간의 관계를 설명하는 데이터 세트가 더 기록된다.

- 각 키워드에 대응하는 각 URL 어드레스로의 중요한 방문(히트)의 누적 회수(여기서는 X 또는 가중치 인자 X 라 부른다). 이것은 각 키워드에 대한 URL의 인기도의 측정값이고 서퍼 행적으로부터 결정된다.
- 초기 소정의 순간에 측정된 중요한 방문의 이전 누적 회수(여기서는 Y 또는 가중치 인자 Y 라 부른다).
- 각각의 상기 웹 페이지의 생성 또는 입력의 순간에 관련된 날짜-시간 인자(여기서는 Z 또는 가중치 인자 Z 라 부른다). Z는 웹 페이지 개발자 검색 엔진에 웹 페이지를 제출하는 날짜-시간이다.

키워드 및 URL 어드레스의 모든 조합이 X, Y, 및 Z에 대한 데이터를 갖지는 않을 것이다.

	키워드	키워드	키워드	키워드	키워드
URL 어드레스 1	X,Y,Z				
URL 어드레스 2					X,Y,Z
URL 어드레스 3			X,Y,Z		
URL 어드레스 4	X,Y,Z				
URL 어드레스 5		X,Y,Z		X,Y,Z	
URL 어드레스 6					
URL 어드레스 7					

테이블 3 정보 제공자(웹 페이지) 및 정보 요청자(키워드) 간의 링크

키워드 URL 링크 테이블을 구비한 신상정보 타입

웹 페이지의 인기도는 서로 다른 그룹의 사람들에 따라 다를 것이다. 다중 신상정보 타입의 포함은 테이블 3의 X, Y 및 Z의 다양한 값을 생성할 것인데, 예를 들면, X1 X2 Y1 Y2 등으로 표시된 글로벌 및 뉴질랜드 인기도 등급을 갖게 될 수도 있다.

	키워드 "스포츠"
럭비와 관련된 URL 어드레스	X1 = 520, X2 = 52
농구와 관련된 URL 어드레스	X1 = 4000 X2 = 20

이 예시에서 럭비 및 농구 URL 어드레스에 대한 (일반적인 신상정보 타입을 이용한) 글로벌 인기도는 520 및 4000 이고, 뉴질랜드 신상정보 타입에서는 각각 52 및 20 이다.

(X1에 기초하여 순위가 매겨진) 일반적인 신상정보 타입 설정이 사용된 경우, 농구 사이트는 최상위에 순위가 매겨질 것이다. (X2에 기초하여 순위가 매겨진) 뉴질랜드 설정이 선택된 경우에는 럭비 사이트가 최고 순위일 것이다. 이것은 뉴질랜드 사람들의 선호도가 반영된 것이다. 이것은 다른 그룹의 사람들에 대한 선호도를 저장하는 매우 간단한 방법이다.

어떤 사람은 뉴질랜드식 럭비 웹사이트가 뉴질랜드 목록상에서 해외 사이트 보다 더 높은 등급을 갖는 것을 기대하겠지만, 이것은 꼭 그렇게 되어야 할 이유는 없다. 스페인의 어떤 사람이 세계에서 가장 멋진 럭비 사이트를 기질 수도 있는 것이다. 상기 시스템은 사용자에게 의한 정보의 인식되어진 품질로만 웹 페이지를 평가한다 - 상기 사이트의 물리적 위치는 중요하지 않다.

매우 간단하게 서로 다른 그룹의 인기도를 획득할 수 있게 하는 다른 국적, 직업, 성별, 나이 등을 나타내는 광대한 범위의 X 값이 있을 수 있다. 사용자는 그들의 개인적인 흥미/성격에 따라 임의의 X 값들을 조합할 것을 선택할 수 있다.

예를 들면,

- X1 은 남성
- X2 은 여성
- X3 은 뉴질랜드 사람
- X4 은 USA
- X1 은 엔지니어
- X1 은 변호사 ...

검색 엔진을 사용하는 '남성' 및 '뉴질랜드 사람'은 X3 및 X1 모두 증가시킨다. 이러한 편익은 상기 시스템의 데이터 요구를 증가시키지만, 서로 다른 사용자들에 대한 검색 결과를 매우 향상시킬 수 있다. 웹 페이지의 총 인기도는 사용자가 하니 이상의 그룹 사람들에게 기여할 수 있는 경우, 개별적인 수로서 저장될 필요가 있다. 개별적인 인기도의 모든 합계는 사용자가 하니 이상의 신상정보 타입에 속할 수 있기 때문에 총 인기도보다 더 크게 된다.

사용자를 위해 시스템을 간단하게 하기 위해서, 옵션을 구비한 디폴트 신상정보 타입(X의 선택)이 특정의 검색을 행하기 위한 다른 신상정보 타입을 사용하는 것이다. 예를 들어, 사용자가 뉴질랜드 남성의 디폴트 신상정보 타입을 기질 수 있지만, 만일 기술적인 검색이 요구된다면 전세계 엔지니어의 누적 검색 지식을 반영하는 '글로벌 엔지니어' 신상정보 타입이 선택될 수 있다.

개인화의 정도는 검색하는 빈도에 의존할 수 있다. 예를 들어, '뉴스' 등의 흔한 키워드는 개인화의 높은 정도(X 값의 큰 범위)를 기질 것이고, '영국 우표' 등의 흔하지 않은 키워드는 개인화기 거의 없거나 이에 없을(오직 글로벌 X 값만 기짐) 것이다. 개인화의 정도는 키워드가 사용된 빈도의 함수(테이블 1로부터 얻어짐)가 될 수 있다.

누적 서퍼 행적 테이블(170)

도 4의 누적 서퍼 행적 테이블(170)의 콘텐츠는 이하에 도시한 테이블 4에서 좀더 상세하게 나타낸다. 테이블 3의 웹 페이지 및 키워드간의 링크에 관한 정보(키워드 URL 링크 테이블(172)로도 불림)는 서퍼 행적 데이터에 의해 업데이트 된다. 상기 누적 서퍼 행적은 모든 개별적인 서퍼 행적으로부터 결합된 정보이고, 각 키워드에 대해 각 웹 페이지가 얼마나 많은 '히트'(중요한 방문)를 기졌니를 결정하는 데 사용된다.

개별적인 서퍼 행적으로부터 수집된 정보는 앞에서 기술된 입력들의 연속이고, 이하에 테이블 타입으로 나타낸다.

IP 번호	사용자 ID	키워드	URL(웹 페이지)	날짜 시간

테이블 4 각 열은 하나의 서퍼 행적 이고 결합된 열은 누적 서퍼 행적임

테이블 3을 업데이트 하기 위해 서퍼 행적 데이터가 처리되는 방법은 이하에 더 설명한다.

신상정보 ID 테이블(166)

도 4의 신상정보 ID 테이블(166)의 콘텐츠는 이하에 도시된 테이블 5에서 좀더 상세하게 나타낸다. 이 테이블은 검색을 수행하기 위해 보통 사용하는 고유 식별, 패스워드, 접속 전자우편, 및 디폴트 신상정보 타입을 포함한다.

사용자 식별	패스워드	전자우편	디폴트 신상정보	또 다른 정보
Joe Bloggs	dogs	jbloggs@AOL	US, 남성	

테이블 5 사용자 식별 테이블

사용자 디폴트 신상정보 타입은 사용자의 개인적인 선호도 신상정보의 부분으로서 저장되는데, 이는 시스템에 소정 타입의 개인적인 식별을 입력함으로써 액세스될 수 있다. 이 정보는 데이터 검색 엔진에 로그인되거나 검색 엔진이 사용자를 식별하기 위해 컴퓨터상에 종래의 기술에서 공지된 '쿠키(cookie)' (로그온 절차와 관련되는 선택 전자우편 주소 및 패스워드(또는 유사한)일 수 있다)를 남길 수 있는 경우에 공급될 수 있다. IP 주소 자체는 개별적인 사용자에게 반드시 유일할 필요는 없기 때문에 충분한 식별 수단은 아니다.

다른 정보는 검색 결과가 어떻게 결합되는지 및 사용자에게 특별한 흥미를 끄는 키워드에 대한 사용자 정의 선호도를 포함할 수 있다. 이 정보는 검색 결과 및 방문하기 위한 웹 페이지의 제사를 능동적으로 개인의 희망에 맞추기 위해 사용될 수 있다.

개인적 링크 테이블(174)

도 4의 개인적 링크 테이블(174)의 콘텐츠는 이하에 도시한 테이블 6에서 좀더 자세하게 나타낸다. 테이블 6은 테이블 3의 구조와 동일하고, 방문 회수 및 키워드를 포함한 각 URL에 관련되어 있는 사용자 개인별 선호도를 기록하기 위해 사용될 수 있다. 그러나, 이 테이블 6에 있어서, Z는 웹 페이지 개별자가 웹 페이지를 제출한 날짜가 아니라 사용자가 웹 페이지를 방문한 날짜-시간이다. 이것은 사용자가 웹 페이지를 방문했던 마지막 시간을 정의함으로써 사용자가 검색을 정제할 수 있게 한다.

	키워드	키워드	키워드	키워드	키워드
URL 어드레스 1	x,y,z				
URL 어드레스 2					x,y,z
URL 어드레스 3			x,y,z		
URL 어드레스 4	x,y,z				
URL 어드레스 5		x,y,z		x,y,z	
URL 어드레스 6					
URL 어드레스 7					

테이블 6 개별 사용자를 위한 정보 제공자(웹 페이지) 및 요청자(키워드)간의 링크

테이블 6의 데이터는 그것을 생성한 개인에 의해서만 액세스되고, 바람직하게는 사용자의 전자우편 또는 IP 주소 변경에 독립적이어서 이러한 변경 중 과거의 개인적인 선호도는 유지될 수 있게 하는 사용자 ID를 사용하여 액세스 가능하다.

이 테이블 6 데이터 세트는 검색 엔진 사이트 또는 개인의 컴퓨터 중 어느 하나에 저장될 수 있다. 로컬 PC상의 저장은 사용자 컴퓨터상에 부가적인 소프트웨어가 설치되는 것이 요구된다. 검색 엔진에 사용자가 일단 실질적인 목표 목록을 갖게 되면 검색 엔진은 웬만하면 변경하지 않고 자주 그곳에 간다는 사실을 포함하는 정보를 저장하는 것은 다수의 이점이 있다.

보안 테이블(168)

도 4의 보안 테이블(168)의 콘텐츠는 이하에 도시한 테이블 7에서 좀더 자세하게 나타낸다. 사용자가 키워드의 인기도를 증가시키기 위해 계속해서 동일한 키워드를 제출하지 않는 것을 보장하기 위하여 다음의 보

인 데이터 테이블이 사용된다. 각각의 엔트리는 정보의 단일 조각, 즉 '에스' 또는 '노'이다. 이 테이블은 키워드와 IP 어드레스 사이의 링크 또는 키워드와 사용자 ID 사이의 링크를 위해 생성될 수 있다.

	키워드 1	키워드 2	키워드 3	키워드 4
IP 어드레스 1	1			
IP 어드레스 2				1
IP 어드레스 3				
IP 어드레스 4				1
IP 어드레스 5		1		

테이블 7 웹 페이지의 인기를 인위적으로 올리기 위해 키워드를 제출하지 않은 컴퓨터 사용자를 확인하기 위한 보안 테이블

앞서 언급한 도 4의 각 테이블을 채우기 위하여 본 발명에 의해 사용된 처리들을 이하에 설명한다.

키워드 테이블(164)을 채우기

이 테이블은 사용자가 검색 엔진에 키워드(52)를 입력할 때마다 채워진다. 제출된 키워드는 테이블 1(키워드 테이블(164))의 키워드 리스트와 비교되고, 아직 존재하지 않는다면 추가된다. 만일 존재한다면, 누적 번호는 하나씩 증가된다. 사용자가 신상정보 타입을 갖는다면, 각 신상정보의 타입에 대한 키워드의 누적 번호 역시 증가될 것이다(W1, W2, W3 등).

웹 페이지 데이터 테이블(URL 테이블)(188)을 채우기

이 테이블은 다음의 여러 방식으로 채워진다.

- 사용자는 테이블 2(URL 테이블(188))에 아직 존재하지 않는 URL 어드레스(126)를 선택한다. URL 어드레스(126) 및 설명(134)은 직접 웹 페이지 데이터 테이블(188)에 가입된다. 신규 URL은 고유식별번호에 할당된다.

- 도 4에 도시된 바와 같이, 단계(176)에서는 웹 페이지 개발자 또는 직접 웹 페이지 데이터 테이블(188)에 가입되는 설명(68) 및 URL(187)을 제출할 수 있다.

- 웹 크롤러는 URL 어드레스 및 설명(웹 페이지의 처음 몇 줄 또는 '제목'이 코드화된 HTML 중 어느 하나의 설명)을 첨가할 수 있다. 이것은 시스템의 중요한 요소는 아니지만, URL 및 설명을 얻는 방법일 수 있다. 이러한 검색 시스템에서, 웹 크롤러는 신규 정보를 찾기 보다는 정보를 증명하기 위해 사용되기 쉽다.

누적 서버 행적 테이블(170)을 채우기

상기 테이블 4에서 참조된 누적 서버 행적 테이블(170)은 '대그된' 웹 페이지가 사용자에게 의해 선택되는 때마다 채워진다. 이것은 서버 행적 정보의 패키지를 보내어, 서버 행적 데이터가 사용자가 웹 페이지 목록으로부터 또다른 웹 페이지를 선택하는 때마다 상기 테이블에 첨가된다.

키워드 URL 링크 테이블(172)을 채우기

누적 서버 행적(170)의 데이터는 키워드 URL 링크 테이블(172)이라 불리는 테이블 3에 기록된 바와 같이 웹 페이지의 인기도(X, Y)를 업데이트 하는데 사용된다. 신규 X 및 Y 값을 얻기 위해 누적 서버 행적(170)으로부터의 데이터를 가지고 테이블 3을 업데이트하는 빈도는 한 시간이, 채워지는 짧은 시간에서부터 한 달이 넘어가는 긴 시간으로의 범위까지 다양하게 변화될 수 있다. 서로 다른 키워드는 다른 기간으로 업데이트 될 수 있다는 점에 유의해야 한다.

누적 서버 행적을 처리하는데 있어서 중간단계는, 누적 서버 히트 테이블을 형성하는 것이다. 이것은 테이블 3의 X, Y의 값을 수정하기 위해 다음에 사용된다.

상술된 바와 같이 키워드 및 URL 사이의 링크('유용한 방문' 또는 '히트')를 기록하는 가장 간단한 방법은 '히트'로서의 서버 행적에서 각 키워드, URL 페어링(URL pairing)을 카운트하게 될 것이다. 좀더 의미있고 세련된 방법은 사용자가 소정의 표준에 접하면 유효로서의 위치 선택된 카운트하는 것이다. 이러한 규범은 한 위치에서 특정 시간을 초과하는 사용자가 될 수 있다. 이 표준이 충족되지 않으면 상기 선택은 테이블 3의 X의 누적값을 증가시키지 못할 것이다.

웹 페이지에서 소비된 시간에 기초하여 X 값을 증가시키는 것도 가능하다. 시간 소비가 많을수록 X 값은 더 증가한다. X는 종 수일 필요는 없다.

로그온 시간, 다운로드 시간, 대역폭, 및 응답 시간이라는 점에서의 웹사이트 성능의 변동으로 인하여, 유효한 '히트'를 나타내기 위해 사용된 소정의 시간은 적절하게 변경될 수 있다. 전문 웹 크롤러가 이러한 데이터를 독립적으로 유효화하도록 사용될 수 있다.

키워드 검색 후 콘텐츠 제공자의 배너의 선택은 그들의 웹 페이지에 대한 히트로서 카운트한다(X 값의 증가). 이것은 그들의 웹 페이지가 키워드와 관련된 인기도 목록에서, 상승하는 것을 가능하게 한다. 이것은 키워드와 함께 보여질 웹 페이지 개발자에게 대가를 지불하게 할 수 있는 기구로서 작용한다. 그들은 인기

도를 높이기 위해 대가를 지불할 수 없다 - 이것은 사람들이 그들의 사이트를 방문하고 거기에서 시간을 소비하며 인기 리스트에 대한 유효한 히트를 기록하는 경우에만 발생할 것이다. 콘텐츠 히트의 값은 사람들이 상기 콘텐츠가 인기도 랭킹에 얼마나 많은 영향을 주는가에 대한 강조에 의거하여 변화(예를 들어, 1 또는 0.5 또는 7)할 수 있다.

이 누적 서퍼 행적 정보는 테이블 8(이하)을 채우기 위한 많은 방법으로 처리될 수 있다. IP 어드레스 또는 사용자 ID 에 따른 누적 서퍼 행적의 그룹화는 개별 사용자를 위한 검색 패턴을 생성한다. 이것은 키워드와 URL 및 시간의 목록이다. 이것은 각 웹 페이지에서 소비된 시간이 각 사용자에게 대해 계산되도록 한다 (상기 웹 페이지로 간 후 기록할 시간이 없으므로 검색기간의 마지막 웹 페이지에서 소비된 시간을 계산하는 것은 불가능하다).

각 방문 사이의 시간이 소정 기간보다 길다면, 1 이 키워드 URL 을 위한 누적 서퍼 히트(α) 테이블에 계산된다(이것은 가장 간단한 방법이고, 예를 들어, 컴퓨터의 검색능력이 사이트에서 소비한 시간에 비례하는 방법 역시 본 발명의 범위 내에 있다).

	키워드	키워드	키워드	키워드
URL 어드레스 1				
URL 어드레스 2	α		α	
URL 어드레스 3		α		α
URL 어드레스 4		α		
URL 어드레스 5				
URL 어드레스 6			α	
URL 어드레스 7				α

테이블 8 누적된 서퍼 행적으로 부터 생성된 누적 서퍼 히트테이블

누적 서퍼 히트는 다음의 방식으로 테이블 3의 X 값을 업데이트 하기 위해 사용된다.

$$X_{(new)} = (X_{(old)} \cdot HF) + \alpha$$

HF 는 0 과 1 사이의 숫자인 히스토리 인자이다. 상기 히스토리 인자는 모든 키워드에 대해 동일할 필요는 없고, 키워드가 사용되는 비율에 따라 변할 수 있다.

테이블 8을 위해 수집된 데이터는 소정의 기간 후에 테이블 3의 X 값을 재계산하기 위해 사용된다. 테이블 3을 업데이트 하는 빈도는 선택된 히스토리 인자(HF)의 값에 영향을 끼칠 것이다. 현존하는 X 에 '히스토리 인자'를 곱하는 이유는 인식된 인기도가 무한히 지속되지 않도록 한다. 상기 히스토리 인자는 과거 인기도에 첨부된 기중을 감소시킨다. 예시으로써 설명하면, 키워드 '스포츠 뉴스'는 (웹 페이지 당 히트수(X)를 기초로 한) 다음의 순위를 갖는 현존의 인기도를 기질 수 있다.

1. 동계올림픽 웹 페이지 $X = 19000$
2. 축구 결과 웹 페이지 $X = 18000$
3. 야구 결과 웹 페이지 $X = 15000$
4. 골프 뉴스 웹 페이지 $X = 15000$

한 주간의 상기 누적 서퍼 히트 테이블은 다음과 같을 수 있다.

1. 동계올림픽 웹 페이지 $\alpha = 500$
2. 축구 결과 웹 페이지 $\alpha = 1800$
3. 야구 결과 웹 페이지 $\alpha = 1500$
4. 골프 뉴스 웹 페이지 $\alpha = 4600$

히트수 변화의 이유는 동계올림픽이 폐회하고 미스터 골프대회가 시작되었다는 사실을 반영한다. 만일 0.9 의 '히스토리 인자'를 갖는다면, 신규 인기도(X)는 다음과 같을 것이다.

1. 골프 뉴스 웹 페이지 $18100(0.9 \times 15000 + 4600)$
2. 축구 결과 웹 페이지 $18000(0.9 \times 18000 + 1800)$
3. 동계올림픽 웹 페이지 $17600(0.9 \times 19000 + 500)$
4. 야구 결과 웹 페이지 $15000(0.9 \times 15000 + 1500)$

따라서, 더 인기있는 웹 페이지가 나타나고 덜 인기있는 웹 페이지는 사라질 수 있고, 이것은 서로 다른 대상 및 사건의 시간차이에 따른 흥미의 변동이 반영된다.

따라서 데이터베이스는 사람들이 인터넷 상에서 무엇이 유용한 것이라고 생각하는지에 대한 강력한 지침을 제공하기 위하여 사람의 마음을 이용한다. 사용자 자신이 그렇지 않으면 이러한 검색을 통해 여과되도록 요구될 실질적인 계산 요청을 대체한다.

테이블 3의 Y 값은 X의 기존값이고, Y 값은 적절하다고 판단되는 간격으로 업데이트 될 것인데, 이 간격은 분, 시간, 일, 주 또는 그 이상이 될 수 있다. 업데이트 간격은 상기 언급한 바와 같이 모든 서로 다른 키워드에 대해 동일할 필요는 없다. 이것은 웹 페이지의 인기도에 대한 변화율을 계산하기 위하여 사용되고 선택 표준으로서 사용될 수 있다.

웹 페이지/URL 링크 테이블의 다른 신상정보 타입

상기 누적 서퍼 행적은 사용자 신상정보에 관한 정보를 포함하여서 테이블 8은 다른 신상정보 타입에 대한 α 의 아래첨자 값을 가지고 계산될 수 있다. $\alpha_1, \alpha_2, \alpha_3$ 등의 이 값들은 X의 아래첨자 값에 대한 신상정보 타입에 대응할 것이다. 이것은 서로 다른 그룹의 사람들의 인기도가 기록되도록 침착한다.

웹 페이지/URL 링크 테이블(172)에 입력되는 신규 웹 페이지 데이터

검색 엔진에 의해 기록된 신규 페이지를 갖는 가장 간단한 방법은 웹 페이지 개발자가 도 4의 조치(176)에 도시된 바와 같은 URL(66), 키워드(70), 사이트 설명(68), 목표 대상(72), 및 날짜-시간(74)을 포함하는 정보를 웹 페이지를 생성 또는 업데이트할 때마다 제출하는 것이다.

이 정보는 직접 테이블 2(도 4의 URL 테이블(188)) 및 테이블 3(도 4의 키워드 URL 링크 테이블(172))을 업데이트 한다. 상기 URL(66) 및 설명(68)은 테이블 2에 입력되고, 페이지가 제출되는 날짜-시간(74)(Z 값)은 각 키워드(70)에 대해 테이블 3에 삽입된다. 사용자는 그들의 웹 페이지를 제출할 수 있도록 하는 키워드(70)의 세트번호를 침착한다. Z 값만이 주어졌을 때 테이블 3이 어떻게 보이는가에 대한 예시가 이 하에 주어진다(형식은 일-월-년).

	키워드	키워드	키워드	키워드	키워드	키워드
URL 어드레스	27/02/98		27/02/98			
URL 어드레스	28/02/98	28/02/98		28/02/98		
URL 어드레스						
URL 어드레스			18/02/98		18/02/98	18/02/98
URL 어드레스						
URL 어드레스				28/02/98		
URL 어드레스		29/02/98				

테이블 9 웹 개발자의 제출자료로부터 생성된 데이터 테이블

테이블 3에서 URL 및 키워드의 조합에 대한 날짜가 없으면, 신규 날짜가 자동적으로 삽입된다. 상기 테이블에 기존 날짜가 이미 존재하면 상기 날짜와 비교되며, 만일 날짜들이 너무 근접하면, 즉 소정의 기간 이하이면, 기존 날짜가 유지되고 신규 날짜는 무시된다. 이것은 사람들로 하여금 웹 페이지를 다시 제출함으로써 신규 웹 페이지 목록의 최상에 올리도록 하는 것을 막는다. 테이블 3의 상기 URL 이 소정의 기간보다 더 근접한 Z 값을 갖는 다른 키워드를 가지는 경우에도, 제출자료가 거부된다. 이것은 웹 페이지 개발자로 하여금 키워드의 다른 세트를 구비한 웹 페이지를 다시 제출하게 하는 것을 막는다.

사용자가 URL을 제출하면, 그들은 테이블 3에서의 사용자들의 특별한 타입(서로 다른 신상정보 타입 Z1, Z2, Z3 등)을 목표로 할 수 있다. 예를 들면, 뉴질랜드 사람들(예를 들면 Z1)에게 특별하게 목표화된 URL 제출자료는 뉴질랜드 사람들이 그 키워드에 대해 검색을 할 때 키워드 신규 목록의 최상에 나타날 것이다. 그것은 어떤 다른 사람이 상기 키워드에 대하여 뉴질랜드 사람들에게 목표화된 URL을 제출할 때까지 최상에 유지될 것이다. 다른 청중에게 목표화된 URL은 뉴질랜드 사람들에게 대한 신규 사이트로서 나타나지 않을 것이며, 또는 대신적으로 그것들은 뉴질랜드 사람들에게 특별하게 목표화된 신규 목록에서만큼 높은 URL로 되지는 않을 것이다.

신규 웹 페이지상의 데이터는 반드시 웹 페이지 개발자에 의해 입력되어야 할 필요는 없다. 그것은 웹 페이지상의 정보가 두드러지게 변경되었을 때마다 검색 엔진에 자동으로 데이터를 제출하는 웹문서 템플릿을 구비함으로써 자동화될 수 있다. 이는 임의의 키워드를 적절하게 변경하도록 웹 페이지 개발자를 촉구할 것이다.

또다른 실시예는 웹사이트 주소 및 키워드를 찾기 위해 전문 크롤러가 파견될 것을 요구하며, 이는 기존 웹 크롤러의 많은 약점을 포함하게 된다. 이러한 실시예는 웹 디자이너가 식별된 키워드를 구비한 페이지를 의도적으로 형성하는 경우에만 효과가 있을 것이다. 또한 이런 일을 행하려는 웹사이트 디자이너/소유자는, 생각컨대 검색 결과상의 즉각적인 목록으로부터 이득을 얻기 위해 검색 엔진에 임의의 업데이트를

기꺼이 제출할 것이다.

이러한 원칙을 연장하면 지동검출에 이르게 되며, 웹 어드레스가 데이터베이스에 키워드 정보를 소유한다면, 그들의 웹 페이지를 쉽게 찾도록 하는 정보를 제공하기 위해 초대장을 지동으로 보낸다. 각 웹 페이지에 대해 제출될 키워드의 이상적인 개수는 기급적이면 50 이하이고, 이미도 대략 5 내지 20의 범위내가 바람직하다. 이것은 또한 웹사이트 디자이너가 그들의 사이트를 설명하기 위해 가장 적절한 키워드를 찾도록 해주며, 그들이 원하는 목표 청중을 선택하게 하는 이점도 있다.

웹 페이지 제출자료 처리는 상기 시스템을 부정으로 사용하는 사람들의 수원을 제한하는 웹 페이지 개별지식별처리를 포함해도 된다. 이것은 다중 웹 페이지 제출자료를 방지하기 위한 보상을 포함할 수도 있다.

신상정보 ID 테이블(166)을 채우기

도 4의 ID 테이블(166)은 사용자로부터 직접 입력됨으로써 채워진다(populated). 사용자가 검색할 때 그들은 계층적인 드롭다운 메뉴로부터 신상정보 타입(54)을 선택할 수 있고, 이것은 다음에 예로 든 것들을 포함할 수 있다.

- 성별(남성 또는 여성)
- 직업(교수, 학생 등)
- 연령 카테고리 등

보여지는 옵션으로부터 사용자가 서로 다른 신상정보 타입을 자신의 디폴트 신상정보 타입으로 저장하고자 하는 경우, 사용자는 이를 선택한다. 이것은 테이블 5(신상정보 ID 테이블(166))에 기록된다. 사용자는 또한 검색 페이지상의 드롭다운 메뉴가 아니라 특정한 개인화 옵션 페이지로부터의 개인화 옵션을 선택할 수도 있다.

개인별 링크 테이블(174)을 채우기

누적 서퍼 행적은 사용자 ID(126)에 의한 정렬에 기초하여 개별 사용자의 검색 패턴을 식별하기 위하여 사용된다. 이 정보는 상기 누적 서퍼 행적(170)이 테이블 3(키워드 URL 링크 테이블(172))을 업데이트 하기 위해 사용되는 것과 동일한 방식으로 개인별 링크 테이블(174)을 업데이트 하기 위해 사용된다. 이 테이블은 사용자 과거 선호도를 지동 북미킹의 타입으로서 저장한다.

보인 테이블(168)을 채우기

사용자가 검색 엔진에 키워드(52)를 매번 입력할 때마다, 키워드(52) 및 IP 어드레스(62)간의 링크를 만듦(또는 키워드(52) 및 사용자 ID(56)간의 링크를 만듦)으로써 보인 테이블(168)을 업데이트 한다. 테이블 7의 데이터는 인기도 목록에 영향을 주는 키워드 검색을 때때로 반복 수행하는 개인을 막는 것 보다는, 인기도 목록(테이블 3의 X 값)에 영향을 주는 체계적인 반복 검색을 막기 위한 목적으로 정기적으로 클리어 된다.

웹 페이지의 목록 결정

테이블 5는 상술된 다양한 데이터 세트를 나타내고, 제시된 웹 페이지의 목록을 다양한 입력과 동작에 의해 제공하며, 이는 이하에 좀더 상세하게 설명될 것이다. 도 5에 도시된 바와 같이, 검색 엔진에 입력된 사용자 데이터는 키워드(52), 사용자 ID(56), 검색 타입(58), IP 어드레스(62), 신상정보 타입(54)을 포함할 수 있다. 또한 어떻게 이 데이터가 웹 페이지(250)의 목록을 결정하기 위해 사용될 수 있으며, 서퍼 행적을 생성하기 위해 어떤 웹 페이지를 태그할 것인가를 결정하는 것(도 3의 단계(118))에 대하여는 이하에 설명한다.

테이블 5의 키워드 URL 링크 테이블(172)에 대응되는 테이블 3의 숫자(X, Y 및 Z)는 다음 타입의 검색들(58)이 주어지기 위해 요청되는 모든 정보를 포함한다.

- 상기 숫자 X 에 기초하여 키워드에 대한 가장 인기있는 URL의 히트목록 순위가 매겨진 인기목록 검색
- 상기 날짜/시간(Z)에 기초하여 키워드에 대한 가장 신규한 URL의 히트목록 순위가 매겨진 핫오프(hot off) 기사 검색
- X 와 Y 간의 차이에 기초하여 가장 최근에 인기가 급상승한 URL의 히트목록 순위가 매겨진 하이플라이 검색
- 히트목록이 상기 숫자 X, Y 또는 Z 중 임의의 것을 갖는 URL의 무작위 샘플인 무작위 검색
- 상기 날짜-시간(Z) 및 특별한 사용자 흥미 날짜(가장 최근의 것만은 아님)에 기초한 히트목록인 창조된 데이터 검색

개인별 링크 테이블(174)은 과거의 선호도도 검색 결과로서 보여질 수 있게 한다..

· 이전 즐겨찾기 검색은 그 개인(테이블 6의 X)에 대한 이전 인기도에 기초하여 순위가 매겨진 히트목록 검색이다. 이 검색은 개인별 사용자의 이전 검색에만 기초한 것이다. 사용자는 그들이 이전에 방문했던 사이트를 매우 빠르게 찾기 위하여 이것을 침착한다.

다른 많은 검색 옵션도 이용할 수 있다.

- 종래의 검색은 정규 검색 엔진을 이용한 목록의 검색이다.
- 다른 콘텐츠만에 대한 검색. 이것은 키워드와 관계되는 광고 등의 다른 콘텐츠의 목록이다.

이러한 검색 결과는 여러가지 다른 방식으로 조합될 수 있다.

- 집합적 검색 순위가 매겨진 히트목록은 상술된 임의의 검색 히트목록의 집합이다(이것은 검색 결과의

디폴트 세트이다).

· 주문형 검색 순위가 매겨진 히트목록은 임의의 상기 목록에 대해 사용자 정의가 된 조합일 수 있다.

도 5는 또한 태그된 웹 페이지를 전송하기 위한 결정(246)에 있어서 키워드 테이블(164) 및 보인 테이블(168)이 사용되는 것을 나타낸다. 이 결정은 키워드 사용의 빈도, 보인 테이블의 데이터 및 사용자 식별의 존재에 기초한 것이다. 태그된 웹 페이지를 여러 사람에게 보내기 위한 상기 결정의 상세한 것은 도 16에 충분히 설명되어 있다.

검색 목록의 다른 타입은 어떻게 구현되는가

이러한 타입의 검색 각각이 어떻게 구현되는가 하는 것은 이하에서 각각에 대한 장점 및 단점들과 함께 자세히 제공된다. 상기 시스템은 사용자의 지적 능력에 의존하고, 이번에는 어떤 종류의 검색을 원하는지를 결정하는 것은 그들이 찾고자 하는 것이 무엇인지에 따라 달라질 것이다. 상기 검색 방법들은 사용자에게 무엇을 사용할 것인가를 알 수 있도록 쉽게 설명되어야 한다.

인기 검색

도 6은 단계(272)에서 키워드(270)의 입력과 연관되는 인기 웹 페이지의 목록을 결정하는 과정을 나타낸다. 이 검색이 선택되고 키워드가 입력되면, 단계(274)가 이어져, 입력된 키워드(270)에 대하여 테이블 3(도 5의 172)으로부터 취한 X 값에 기초하여 웹 페이지의 목록을 생성한다. 이러한 웹 페이지는 테이블 3의 고유 웹 페이지(URL) 번호에 의해 식별된다. 따라서, 단계(274)로부터 얻어진 웹 페이지 번호의 목록은 단계(276)에서 URL 어드레스 및 테이블 2(180, 도 5)의 웹 페이지 설명과 결합된다. 웹 페이지의 결과 목록은 상술된 도 5의 단계(246)의 결과에 의거하여 단계(278)에서 태그된 후, 사용자의 선택을 위하여 사용자에게 보내진다.

핫오프 기사 검색

도 7은 단계(290)에서 입력된 키워드와 관계있는 신규 웹 페이지의 목록을 결정하는 과정을 나타낸다. 이 검색이 선택되고 키워드가 입력되면, 단계(294)가 이어져, 단계(290)에서 입력된 키워드에 대하여 테이블 3(도 5의 키워드 URL 링크 테이블(172))로부터 취한 Z 값에 기초하여 웹 페이지의 목록을 생성한다. 이러한 웹 페이지는 테이블 3의 고유 웹 페이지(URL) 번호에 의해 식별된다. 따라서, 단계(294)로부터 검출된 웹 페이지 번호의 목록은 단계(296)에서 URL 어드레스 및 테이블 2(도 5의 URL 테이블(188))의 웹 페이지 설명과 결합된다. 웹 페이지의 결과 목록은 상술된 도 5의 단계(246)의 결과에 의거하여 단계(298)에서 태그된 후, 사용자의 선택을 위하여 사용자에게 보내진다.

사용자는 각 웹 페이지가 정확히 제출된 시점을 알 수 있으므로, 인터넷 사용자는 그 흐름을 알 수 있게 된다. 이러한 특징의 간접적인 결과는 웹 디자이너로 하여금 그들의 사이트를 업데이트 하는 동기를 부여하는 것이다. 신규 및 업데이트된 사이트에 주어진 저명성은 인기 히트목록상에 확고한 입지를 갖게 하는 수단을 제공하고, 적절한 키워드의 사용을 장려하며, 사용자에게 유용하다고 판단하는 웹 페이지들이 계속 유지될 수 있도록 한다.

하이플라이 검색

도 8은 단계(320)에서 입력된 키워드와 연관되는 하이플라이 웹 페이지 검색을 나타낸다. 이것은 웹 페이지의 인기도가 가장 빠르게 증가하는 목록에 대한 검색이다. 이 검색이 선택되고 키워드가 입력되면, 단계(324)가 이어져, 단계(320)에서 입력된 키워드에 대하여 테이블 3(172, 도 5)로부터 취한 X 및 Y 값 사이의 관계에 기초하여 웹 페이지의 목록을 생성한다. 이러한 웹 페이지는 테이블 3의 고유 웹 페이지(URL) 번호에 의해 식별된다. 따라서, 단계(324)로부터 검출된 웹 페이지 번호의 목록은 단계(326)에서 URL 어드레스 및 테이블 2(188, 도 5)의 웹 페이지 설명과 결합된다. 웹 페이지의 결과 목록은 상술된 도 5의 단계(246)의 결과에 의거하여 단계(328)에서 태그된 후, 사용자의 선택을 위하여 사용자에게 보내진다.

하이플라이 목록은 테이블 3의 기존 인기순위(Y) 및 신규 인기순위(X)를 비교하여 계산된다. 이것으로부터 히트의 백분율 증가가 계산된다. 대안적인 방법은 지난번과 비교하여 상승된 장소의 번호에 의해 인기도의 변화율의 순위가 매겨지는 것이다.

본 실시예에 대한 인기도의 변화율 계산 공식은 다음과 같다.

$$((X - Y) / Y) \cdot (X / (X_m \cdot \beta))$$

여기서 X_m 은 키워드에 대응하는 X의 최대값이고, β 는 인기도 목록의 최상 및 최하에서의 변화의 상대적인 중요도를 바꾸기 위해 변경될 수 있는 부가 변수이다.

X의 최대값을 공하는 이유는 낮은 인기도 값에서의 작은 변화가 테이블의 상부에서의 더 중요한 변화를 제거하지 않게 하는 것을 보장하기 위해서이다. 예를 들면, 다른 웹 페이지는 낮은 퍼센트 증기를 나타내지는 않을지라도 520 히트에서 4000 히트(더 많은 상당한 증가)까지의 증기를 나타낼지도 모르는 반면에, 미리 단일 선택만이 기록된 것을 구비하고 그 다음날 5 히트를 끌어내는 웹사이트는 500% 증가한 퍼센트를 나타내게 되는 것이다.

무작위 검색

이것은 입력된 키워드와 관계있는 임의의 X, Y, 및 Z 값을 갖는 웹 페이지의 무작위 선택에 기초하여, 통상의 경우를 벗어난 웹 페이지를 보기 원하는 사용자를 위하여 덜 인기있는 웹 페이지를 무작위로 선택하는 것을 말한다. 따라서, 사용자가 도 9에 도시된 단계(352)에서 키워드를 입력한 후, 도 5에 도시된 키워드 URL 링크 테이블(172)이 참조되고, 단계(354)에서 도시된 비와 같이 무작위 번호 생성기를 사용하여 지동으로 생성된 웹 페이지 번호의 무작위 목록이 결정된다. 이러한 무작위 선택에 있어서는 키워드와 관계있는 X, Y, 또는 Z 값을 갖는 웹 페이지만 선택되는데, 이는 과거의 어떤 단계에서 사용되었거나, 혹은 웹 페이지 개발자가 자신의 웹 페이지가 키워드에 어느 정도 관계가 있다고 판단하였음을 나타내기 때문이다.

그 후, 단계(354)로부터 검출된 웹 페이지 번호의 목록은 단계(356)에서 URL 주소 및 테이블 2(188, 도 5)의 웹 페이지 설명과 결합된다. 웹 페이지의 결과 목록은, 상술된 바와 같이 도 5의 단계(246)의 결과에 의거하여 단계(358)에서 태그된 후 선택을 위해 사용자에게 보내지게 된다.

종래의 검색

도 3에서 다른 검색 엔진(116)으로 언급된 종래의 검색 엔진의 일반 검색방법으로서, 이것은 전술한 바와 같이 사용자의 옵션으로서, 본 발명에 따른 검색과 함께 포함될 수도 있고 포함되지 않을 수도 있다.

컨텐츠만 검색

이것은 키워드와 연관된 광고 등과 같이 사용자들이 제어할 수 없는 컨텐츠의 목록이다. 위에서 설명되는 바와 같이, 본 발명의 바람직한 실시예에 따르면 가장 많이 지불한 것들이 목록의 최상부에 있을 것이다. 물론, 컨텐츠 제공자들에게 지불하는 순서를 식별하는 다른 시스템도 역시 구현될 수 있다.

이전 즐겨찾기 검색

도 10은 개인사용자의 이전 검색에만 기초한, 이전 즐겨찾기 검색을 도시한 것이다. 이는 사용자들이 이전에 방문했던 사이트를 매우 빠르게 찾을 수 있도록 하며, 따라서 지동 북마킹을 실행한다. 바람직하게는 본 발명에 따른 검색 엔진 시스템에 로그인하기 위해 패스워드가 사용되기 때문에, 사용자는 임의의 컴퓨터로부터도 그들의 개인별 즐겨찾기를 액세스할 수 있다.

따라서, 도 10에 나타난 바와 같이 사용자가 단계(372)에서 키워드를 타이프하면, 도 5에 도시된 개인 링크 테이블(174)로부터 상기 키워드로 즐겨찾는 사이트(이전 사용에 기초한)가 무엇인지를 결정하는 단계(374)가 뒤따른다. 사용자는 시스템에 로그인하는데 사용되는 패스워드를 가지고 있기 때문에, 사용자는 어떠한 컴퓨터로부터도 개인별 즐겨찾기를 액세스할 수 있다.

따라서, 이러한 검색성능으로 인해, 수동으로 웹 페이지를 북마크할 필요가 없다. 만약 사용자가 예를들면 'marbles' 같은 좋은 사이트를 북마크하는 것을 잊어버렸다면, 그 사이트로 갈 수 있는 키워드를 다시 타이핑함으로써 쉽게 찾을 수 있다. 사용자 즐겨찾기가 변경되면 이는 개인 링크 테이블(174)에 반영될 것이다.

개인별 즐겨찾기 검색의 또다른 실시예는, 키워드를 사용하거나 또는 사용하지 않고, 웹 페이지를 최근에 방문했던 날짜를 명기하는 것을 포함한다. 그런 후 웹 페이지는 도 5의 개인 링크 테이블(174)의 Z에 기초하여 순위가 매겨진다. 예를들면 만약 사용자가 작년의 중반에 한 사이트를 찾았다면, 사용자는 그 날짜로 검색을 정제할 수 있으므로, 비록 사용자가 관련 키워드를 기억하지 못하더라도 이전의 유용한 웹 페이지를 좀 더 쉽게 찾을 수 있도록 한다.

이러한 지동 북마킹 특성은 특정의 컴퓨터에 의해 시도되는 인터넷 사용의 유형을 모니터링하는 도구로 적용될 수도 있으며, 따라서 예를 들면, 성인 웹 페이지 등과 같은 바람직하지 못한 사이트에 액세스하는 것에 대해 아이들/피고용자의 부모/고용주에게 경고를 제공할 수 있다. 바람직한 실시예에서는, 스스로 컴퓨터를 사용할 것 같지 않은 부모/고용주들을 위해서, 이러한 사용의 통지가 아이들/피고용자들이 방문한 웹 페이지 및 선택된 키워드의 목록이 기록된 편지로서 부모/고용주에게 자동으로 제공된다. 이러한 정보는 도 5의 각 사용자 테이블로부터 직접 찾아진다. 이것은 부모/고용인의 정보를 포함하는 사용자 식별 코드를 요구한다.

집합 검색(collective search)

도 11에 도시된 집합 검색은 본 발명에 따른 디폴트 검색이며, 사용자가 다른 검색 옵션을 선택을 하지 않았을 때 사용된다.

단계(402)에서 키워드를 입력하면, 상기 키워드는 그 키워드와 관련된 웹 페이지 무음의 조합으로부터의 선택을 위해 이용된다. 예를 들면 단계(404)에 도시된 바와 같이, 종래검색, 인기검색, 하이플라이 검색, 신규검색 및 과거선택검색 결과에 대하여 동일한 기준치를 둔 조합이 웹 페이지 수의 목록을 얻는데 이용된다. 그런 다음에, 단계(406)에서는 단계(404)로부터 발견된 웹 페이지수의 목록이 URL 주소 및 테이블 2로부터의 웹 페이지 설명(도 5의 188)과 결합된다. 그 후 단계(408)에서 웹 페이지의 결과 목록은 상기된 것처럼 도 5의 단계(246)의 결과에 따라 태그되고, 사용자가 선택할 수 있도록 전달된다. 처음에 시스템이 구성되었을 때, 검색 엔진(10) 데이터베이스는 인기있는, 하이플라이 및 신규 웹 페이지 히트목록에 대한 어떠한 정보도 가지고 있지 않을 것이며, 따라서 검색결과는 처음에는 종래의 히트목록(통상적 검색 엔진)으로부터 얻어질 것이며, 그 후에는 위에서 설명한 바와 같이 태그된 웹 페이지가 데이터베이스 세트를 생성하는데 사용된다. 시스템이 발전함에 따라, 각각의 다른 검색 방식과 관련된 데이터 세트가 자리집개(populate) 될 것이고, 다른 검색 방식을 사용하는 검색이 점점 더 유용해 질 것이다.

날짜 생성 검색

도 12는 웹 페이지가 제출된 날짜를 사용자들이 선택하도록 하는 날짜 생성 검색을 도시한다. 이러한 특성은 도 4에서 날짜와 시간 제시(74)로서 규정된 날짜 생성 데이터 엔트리를 포함한 웹 페이지에 대해서만 적용할 것이다. 단계(432)에서 날짜-시간 및/또는 키워드가 입력되지만, 검색 엔진(10)은 이러한 변수들에 연관된 웹 페이지 수의 목록이 얻어지는 단계(434)를 수행할 것이다. 그런 후에, 단계(436)에서는 단계(404)에서 발견된 웹 페이지의 수의 목록이 테이블 2에서의 URL 주소 및 웹 페이지 설명(도 5의 188)과 결합된다. 그 후 단계(438)에서는 웹 페이지의 결과목록이 상기한 바와 같이 도 5의 단계(246)의 결과에 따라 태그되고, 선택을 위해 사용자에게 보내진다.

맞춤 검색

도 13은 사용자들이 그들의 디폴트 히트목록이 어떻게 나타나기를 원하는지를 결정하도록 하는 맞춤 검색이다. 단계(462)에서, 맞춤 검색을 개시하기 위해 키워드 및 사용자 ID가 선택된다. 하지만, 도 11에 관련하여 기술한 집합 검색의 단계(404)와 동일한 단계(466)에서 맞춤 검색을 개시하기 이전에, 히트목록의 사용자 디폴트 혼합을 사용자에게 맞추기 위한 단계(464)가 적용된다. 예를 들면 사용자는 디폴트 검색 결과

가 하이퍼링크 웹 페이지가 아닌 단지 인기있고 새로운 웹 페이지인 것만 포함할 것을 원할 수 있다. 그런 다음에는 웹 페이지 수의 목록을 생성하는 맞춤형 검색이 단계(466)에서 수행된다. 그 후에는, 단계(466)에서 발견된 웹 페이지수의 목록이 단계(468)에서 테이블 2(도 5의 188)에서의 URL 어드레스 및 웹 페이지 설명으로 조합된다. 도 5의 단계(246)의 결과에 따른 웹 페이지의 결과목록은 단계(470)에서 태그된 후, 선택되도록 사용자에게 보내진다. 바람직한 일실사에서는, 사용자 자신의 검색 패턴을 기초로 한 변형 맞춤형 검색을 생성하기 위해, 디폴트 검색 결과목록의 구성은 사용자의 행위로부터의 '학습'에 의해 수정되어질 수 있다. 예를 들어 사용자가 새로운 웹 페이지 또는 하이퍼링크 웹 페이지를 일관되게 선택한다면, 디폴트 검색결과와 설정은 정규적인 검색스타일을 반영하여 변경될 것이다.

매거진 검색

본 발명에 따른 매거진 검색은 특정 키워드를 입력하는 대신에 사용자로 하여금 일련의 메뉴구동형 객체선택(또는 유사한 계층 구조)에 의해 검색을 할 수 있도록 한다.

기존의 매거진형 검색 엔진은 사용자/뷰어에게 전송되는 주어진 페이지상의 공간에 대해 편집자로 하여금 정보의 구조를 셋팅하고, 적절한 가치를 결정하며 가격 등의 평가기준을 마련할 것을 요구한다. 본 발명의 검색시스템을 사용함으로써, 사용자 자신이 불만한 가치가 있는 것과 없는 것을 동적으로 결정한다. 따라서, 객체의 계층에 대하여 편집에 관한 입력을 필요로 하지는 아니, 이러한 객체 각각에 대해 가장 인기있는 것으로 나타나는 웹 페이지는 자동적으로 발전할 것이다.

서로 다른 그룹의 사람들에 대한 데이터 세트 사용

다양한 문화, 지역, 직업, 성별, 또는 연령별 관심을 반영한 결과를 제공하기 위해 다양한 인기 히트목록이 사용될 수 있다. 따라서, 도 14에 도시된 바와 같이, 어떤 사용자가 단계(490)에서 키워드 및 사용자 ID를 입력하면, 그 사용자의 신상정보와 동일한 '집단'의 사람들이 보기를 희망하는 웹 페이지의 유형을 반영하기 위하여 상기 사용자의 디폴트 신상정보가 사용될 수 있다. 따라서, 단계(494)에서 행해지는 검색은 단계(492)에서 얻어진 사용자 개인에 관한 신상정보와 유사한 '집단' 소속의 사람들의 디폴트 신상정보로부터 얻어진, 상기 X, Y, Z로 붙여진 값들에 기초한다. 전반적으로 광범위한 검색결과보다는, 검색결과가 사용자와 동일한 집단에 대해 특정화되어 얻어진다. 상기한 바와 같이, 단계(496) 및 단계(498)로부터 이끌어낸 웹 페이지의 결과 목록은 상기 집단에 대해 특정화된다.

따라서, 예를 들면, 지역적요소로서 선택되는 뉴질랜드라는 신상정보 타입을 가진 특정 사용자에게 대한 팀 이외스포츠에 대한 검색은 럭비에 대한 자료가 두드러지게 나타날 수 있는 반면에 미국의 신상정보 타입은 야구/미식축구 자료에 대하여 편입될 수 있다. 특정 신상정보 타입에 연관된 인기 히트목록의 문맥에 따라, 이러한 기법은 같은 단어의 다양한 의미를 구별해내는 능력을 제공한다. 예를 들면 'accommodation'이라는 키워드를 사용한 일반적인 검색은 주택, 임대 등에 연관된 결과를 포함하는 반면에, 사용자의 신상정보 타입에 검인(optometry)에 대한 관심이 나타나 있는 사용자라면, 'accommodation'이라는 용어는 상당히 다르게 해석될 것이다.

검색 엔진 작업지 또는 사용자에게 의한 사이트에 대한 어떠한 능동적인 평가도 없이, 사이트의 관련성은 자동적으로 발전할 것이다. 사용자의 특정 타입에 대한 웹사이트 관련성을 분석하기 위해 요구되는 복잡한 알고리즘이 없다. 대신에, 관련이 있다고 여겨지는 사이트의 유형은 그들의 신상정보 유형의 특성을 선택하는 사용자 예를 들어 임박등반에 관심있는 미국여성에 의해 결정될 것이다. 관련성이 더욱 많은 사이트는 좀더 높은 히트를 이끌어 낼 것이며, 그들의 순위가 높아지면서 다음 사용자가 그 사이트를 조사할 가능성 또한 증가될 것이다. 상기 예에 있어서, 검인, 시아, 렌즈, 시력 등과 관련이 없는 'accommodation'이라는 키워드에 대해 목록화된 웹사이트는 적당한 히트를 나타내는데 요구되는 기간동안은 액세스되지 않을 것이다. 따라서, 매우 낮은 순위를 받을 것이며, 이로 인해 심지어 다른 사용자에게 의한 액세스도 저조한 것처럼 보일 것이다.

사용자는 단일 세션동안 다양한 검색에 대한 다양한 신상정보 타입을 선택할 수 있으며 디폴트 신상정보 타입에 제한받지 않을 것이다.

본 발명의 또 다른 일실사에 있어서, 어떤 그룹의 사람들에 있어서 그들의 검색 결과가 그 그룹과 관계 있는 데이터베이스를 업데이트하는 목적을 위해 실제로 사용되는 사람에 대한 증명의 수준이 포함될 수 있다. 예를 들면, 의사로서 식별된 사용자 ID를 가진 의사가 소정의 의학적 상태에 관련된 검색을 수행할 수 있으며, 상기한 바와 같이 그러한 선택은 태그되어 의사의 집단에 대한 데이터베이스에 사용될 수 있다. 하지만, 환자가 의사와 같은 집단의 검색으로 환자의 신상정보를 식별받기 원하더라도, 그들의 선택은 진짜 의사의 선택만큼 중요하진 않으므로, 의사들에게 가장 적절하게 보이는 웹 페이지 목록을 볼 수 있는 반면에, 그들의 ID가 의사로서 식별되지 못했기 때문에, 그들의 선택이 의사집단의 데이터베이스를 업데이트하는 데는 사용되지 못한다.

검색 옵션을 제한하는 방법

도 15에 기술된 본 발명의 또 다른 특성은, 키워드 배제기(keyword eliminator) 특성이 작동중일 때에는 어린이 등의 소정의 사용자가 바람직하지 못한 키워드 및 웹 페이지를 검색하는 것을 막는 키워드 배제기의 특성이다. 본 발명자가 인지한 바에 의하면 예를 들어, 외설적인 콘텐츠가 있는 인터넷상의 모든 사이트로의 액세스를 추적하고 막는 것을 시도하는 것보다는 포르노에 대한 아이들의 검색을 방지하는 것이 잠재적으로 훨씬 용이하다. 이것은 기존의 '망지기(net nanny)'형 장치에 대한 보조적인 도구로서 사용될 수 있다. 따라서, 도 15에 도시되는 바와 같이, 키워드 배제기가 작동중이면, 기존의 테이블 액세스 불가 키워드가 테이블에 저장되고, 단계(520)에 도시된 바와 같이, 키워드가 먼저 입력되면 단계(522)에서 비교된다. 따라서, 액세스 불가 키워드는 검색되지 않을 것이다. 따라서, 예를 들면, 부모들은 아이들이 검색하는 것을 원치않는 키워드(552)의 타입을 선택할 수 있다. 그리고 이것은 부모들이다 다룰 것이다. 시스템은 단계(524)에서 다음 검색에 사용될 수 있는 키워드를 골라서 이를 배제한다.

어떤 사용자를 표본으로 할 것인지를 결정하는 방법

도 16은 어떤 검색 결과가 샘플링되고, 또한 도 4의 누적 서퍼 행적을 나타낸 테이블 (170)(테이블 수로도

지칭됨)을 구성하는데 사용될 것인지를 결정하는 과정을 나타낸다. 기능하기는 하지만, 모든 검색 히스토리
 니에 관련되어 데이터를 수집할 필요는 없으며, 이것은 결과의 어떠한 세트가 '태그된' 웹 페이지로 배달
 되게 할 것인지를 결정하는 것에 의해 제어될 수 있다. 이것에 관하여는 앞에서 예시로서 제시된, 의사 같
 은 특정 그룹에 대한 식별된 사용자의 검색에 관하여 이미 언급되었다.

단계(556)에 의해 도시된 비외 같이, 사용자 ID를 가지며 상기 도 5의 개인 링크 테이블 (174)을 이용할
 것을 선택한 사용자에게 대해서, 개인 링크 테이블 (174)에 그들의 과거 모든 취향을 기록하기 위해서 그들
 의 결과를 '태그' 할 필요가 있다. 본 발명 시스템에 따른 검색 엔진은 소정의 보안 수준을 만족시키지 못
 하면 (이에 대해서는 아래를 참조) 사용자의 개인별 취향은 업데이트할 수 있지만 테이블 3은 업데이트할
 수는 없다. 하지만, 개인 링크 테이블 (174)이 중앙 컴퓨터가 아니라 개인 컴퓨터상에 저장되어 있다면,
 데이터는 국부적으로 저장되므로 태그된 결과를 보낼 필요가 없다.

단계(558)에 의해 도시된 비외 같이, 키워드가 제거되면, 사용자에게 태그된 결과의 세트가 전송되기 전에
 보안 테이블 (168) (테이블 7)을 이용해서 IP 어드레스(62)가 이미 키워드를 검색한 적이 없는지를 체크하
 게 된다. 만약 검색한 적이 있다면, 비록 누적 서퍼 행적 (170) (테이블 4)에 제공되지는 않지만 사용자는
 검색을 계속 수행할 수 있다. 이것은 모든 정상적인 사용자들이 인기 히트목록에 영향을 미칠 수 있도록
 하며 모든 사용자들이 그들이 원하는 것이 무엇이건 간에 검색할 수 있도록 하면서도, 스팸머(spammer) 등
 시기꾼 같은 사용자가 인기 히트목록에 영향을 미치지 못하게 한다. 보안 테이블(168)은 반복되는 검색을
 검출해내기 위해서 키워드(52)와 사용자 ID간의 링크에 대한 정보를 또한 포함한다.

사용자가 자신의 컴퓨터의 IP 어드레스를 변경하는 것이 기능하기는 하지만, 이것은 IP 넘버의 사용을 등
 록하고 추적하는 수많은 방법을 통해서 검출될 수도 차단될 수도 있다.

잘못된 검색을 차단하는 다른 방법으로는:

- 검색 엔진으로 기록된 사용자 ID (554)를 갖는 사용자에게 대해서만 서퍼 행적을 생성시키는 방법
- 방문이 유용한 히트로서 카운트되는데 요구되는 시간 제한을 연장시키는 방법
- 키워드로부터 URL에 단일 방문은 카운트하지 않는 방법 (경고 시간을 측정할 아무런 수단도 없기
 때문에)

을 포함한다.

단계(560)에 의해 도시된 비외 같이, 인기 키워드는 10번에 1번, 100번에 1번, 심지어 1000번에 1번의 발
 생조치 추적될 수 있으며, 이러한 선택의 빈도는 시스템을 최적화시키기 위해 변경될 수 있다. 키워드 사
 용의 빈도는 도 5(테이블 1)에 도시된 비외 같은 키워드 테이블(164)로부터 결정된다. 태그된 결과의 전송
 빈도도 또한 다양한 키워드에 대한 인기도가 변경되는 속도에 링크될 수 있다. 예를 들면, 'IBM'이라는 키
 워드는 이미 최상위에 IBM의 홈페이지를 가질 것이고 대부분의 사용자들은 그 홈페이지로 갈 것이지만,
 '최신 유행(latest fads)'라는 키워드는 더욱 자주 샘플화될 필요가 있는 계속 변경되는 웹 페이지의 세트
 를 가질 수 있다.

도 5(테이블 3)의 키워드 URL 링크 테이블이 지나치게 커지는 것을 피하기 위한 하나의 방법은, 키워드들
 이 소정의 사용빈도에 도달했을 때만 테이블 3에 등록시키는 것이다. 이것은 사용된 빈도가 적은 키워드
 (테이블 1로부터 찾아)를 태그된 결과로 전송하지 않도록 함으로써 제어된다.

방문할 웹 페이지에 대한 능동적 제한

본 발명의 또 다른 특성은 도 17에 도시되며, 웹 페이지를 능동적으로 제한하기 위해 데이터를 사용하는
 것과 관련된다. 이것은 사용자가 요청한 다음, 검색주제(subject)에 대한 새로운 데이터가 존재하면 사용
 자가 이를 통지 받기 때문에 검색과는 다르다. 이것을 수행하기 위해서는 사용자가 자신이 흥미 있어 하는
 키워드 및 사용자가 필터 또는 매개물로서 적용되기를 원하는 신상정보 타입 및 검색 타입(새로운, 히이플
 리아, 인기있는) 중 어느 것을 능동적으로 지정해야만 한다. 이러한 정보는 도 5 (테이블 5)에 도시된 사
 용자 ID (166)에 저장된다.

따라서, 사용자는 비슷한 생각을 갖는 사람들의 집단에 의해 결정되어 제시된 웹 페이지의 목록을 다양한
 간격으로 받는다. 예를 들면, 사용자는 다음과 같은 웹 페이지를 통지받기를 원할 수 있다.

키워드(582)	신상정보 타입(매개물588), 검색 타입(586)
럭비	뉴질랜드, 남성, 하이 플라이
충치치료	치과 의사, 신규

이러한 방식으로, 다른 뉴질랜드 남성들이 유용하다고 찾아낸 '럭비'에 대한 하이플라이 웹 페이지가 존재
 한다면(즉, 그들은 이 정보를 찾는 데 막대한 시간을 허비하였다. - 테이블 3의 X의 높은 변화율) 사용자가
 이를 통지받게 된다. 이와 유사하게, 만약 치과 의사가 찾는 '충치치료'에 관련된 어떤 새로운 정보가 제
 시된 것이 있다면, 사용자는 이를 알게된다(테이블 3의 Z값). 컴퓨터 매개물이 선택된 동료집단만큼 정보
 를 항상 잘 걸러줄 가능성은 별로 없다. 다른 '매개물 타입' 소프트웨어에 대한 본 시스템의 장점은 사용
 자의 컴퓨터상에 아무런 소프트웨어를 필요로 하지 않는다는 것이다. 다른 검색 엔진 데이터 세트에는 당
 연한 확장으로 모두 포함된다.

제안된 웹사이트는 사용자가 다음에 검색 엔진을 액세스 할 때 또는 사용자가 이메일(e-mail) 통지를 통해
 서 상기 제안된 웹 페이지를 통지 받기를 원할 때 사용자에게 보여질 수 있다. 이러한 방식의 웹 페이지는

키워드에 대한 어떠한 능동적인 검색 없이도 사용자의 관심을 끌 수 있다.

빙문할 웹 페이지에 대한 수동적 제안

본 발명의 또 다른 특성은 도 18에 의해 기술되며, 사용자로부터 아무런 능동적인 입력을 요구하지 않으며 사용자의 과거에 검색했던 방법에 기초하여 자동적으로 웹 페이지를 제안하는 것과 관련된다.

단계(620)에 도시된 바와 같이, 사용자 ID가 입력되는 순간, 시스템은 다양한 간격 또는 시간(검색 엔진에 로그인하는 때 등)에, 사용자가 도 5(테이블 6)의 테이블(174)을 이용해서 어떤 키워드, 신상정보 타입 및 검색 타입을 빈번하게 찾았는지를 찾음으로써 자동적으로 작동될 수 있다. 예를 들면, 사용자가 '뉴질랜드', '남성'으로서 학비정보를 빈번하게 찾고 '치과의사'로서 충치치료를 찾을 수 있다. 이러한 정보는 앞에서 개인 링크 테이블(174)이라 언급된, 자동 북 마킹 테이블로부터 찾아낼 수 있다. 만약 사용자가 소정의 시간동안 이러한 객체를 찾아내지 못하고 있고, 또한 신규 또는 하이 플라이 정보 소스가 존재한다면, 사용자는 신규 정보 소스를 자동적으로 통지 받을 것이다.

본 발명의 수정예에 있어서, 사용자의 키워드에 관련되어 두개의 최신 및 최고 하이플라이 사이트에 대한 주기적인 이메일이 보내질 수 있다.

제안된 키워드의 목록 결정

많은 사용지들에 있어서 인터넷 검색의 문제점은 검색을 위해 어떤 키워드를 사용할 것인가이고 알려져 있다. 본 발명은 막대한 수의 키워드를 제공으로 구현될 수 있기는 하지만, 사용자들이 선택할 키워드(여구 포함)의 수가 너무 많은 것은 문제가 될 수 있다.

따라서, 도 19에 도시된 바와 같이, 본 발명은 단계(640)에서 소정의 신상정보 타입에 따라 입력된 키워드에 대한 동의어를 제공하는 데이터 세트(642)를 또한 구비한다. 도 19에 나타나 있는 시스템은 키워드 제 시지(keyword suggester)라 부르기로 한다. 일상시에서, 이것은 단계(640)에서 사용자에 의해 입력된 키워드를, 다른 검색방법에 의해 다른 사용자에 의해 시도되었던 도 5(테이블 1)의 키워드 테이블(164)에 존재하는 기존의 키워드 및 어구의 단계(646)에서 매칭시킴으로써 제공된다. 각 키워드는 단계(660)에서 태 그된 후, 단계(662)에서 사용지에 의해 선택된 키워드는, 상기 웹 페이지 서버 행적으로서, 사용자가 입력한 원래 키워드, 선택된 키워드(652), IP 어드레스(130), 사용자 ID(128) 및 날짜-시간(132) 데이터를 포함하는, 도 19에 도시된 키워드 서버 행적(648)을 만드는 데 사용된다.

이후 누적 키워드 서버 행적(648)으로부터의 데이터는 키워드들 간의 링크를 강화하는데 사용된다. 이러한 방식으로 시스템은 어떤 키워드들이 서로 연관되는 지를 학습한다. 시스템은 어떤 URL이 키워드와 관련이 있는지를 학습하는 것과 같은 방식으로 어떤 단어가 서로 연관된 것인지를 학습한다. 제안된 키워드의 목록은 키워드 제안자가 사용될 때마다 관련성이 향상되므로 시간이 지니면 좀 더 관련성이 높게 될 것이다.

제안된 키워드를 결정하는 데이터 세트를 생성하는 방법

도 20에 도시된 바와 같이, 키워드 링크 테이블(696) 및 누적 키워드 행적 테이블(698)은 제안된 키워드에 대한 데이터 세트를 생성하기 위하여 상기 보안 테이블(168)과 함께 사용된다. 아래 테이블 10에 나타난 키워드 링크 테이블(696)은 제안된 키워드 목록으로부터 얼마나 자주 각 키워드가 선택되었는지를 기록한다. 이후 이것은 서로 다른 키워드들의 유용성에 대해 서로 순위를 매기는데 이용된다.

	키워드 1	키워드 2	키워드 3	키워드 4	키워드 5
키워드 1	-	5			
키워드 2	20	-	1134		
키워드 3		356	-		
키워드 4				-	
키워드 5		20			-
키워드 6			3		
키워드 7					168

테이블 10 키워드 링크 테이블

테이블 10에서 키워드(2)를 입력한 사람은 키워드(3)을 가장 유용한 것으로 찾았으며 그 다음으로는 키워드(5)를, 그 다음은 키워드(1)을 찾은 것을 알 수 있다. 키워드는 지향성(directional aspect) 기지는데, 예를 들면, 키워드(3)는 키워드(2)를 시도한 이후 1134번 유용한 것으로 찾았었다. 하지만 키워드(2)는 사용자들이 키워드(3)을 시도한 이후 고작 356번 찾았었다.

테이블 10의 키워드들 간의 링크에 관한 정보는, 제안된 키워드(키워드 서버 행적(648))를 사람들이 사용

하는 방법에 대한 정보에 의해 업데이트된다. 누적 키워드 서퍼 행적(698)은 모든 개개의 키워드 서퍼 행적(648)로부터의 조합된 정보이며, 각 키워드가 각 키워드에 대해 얼마나 많은 '히트'(중요한 방문)를 기록하는지를 결정하는데 이용된다.

각 개개의 서퍼 행적으로부터 수집된 정보는 아래 테이블 11의 테이블 형식에 나타난 누적 키워드 서퍼 행적이 되는 일련의 입력이다.

IP 번호	사용자 ID	(최초) 키워드	(제시된) 키워드	날짜-시간

테이블 11 키워드 누적 서퍼 행적

키워드 링크 테이블을 증가 시키는 방법

도 20도 역시 검색 엔진에 입력되는 사용자들의 일련의 키워드를 기록함으로써 테이블 10의 키워드들 간의 링크가 개시될 수 있는 방법을 도시한다. 만약, 예를 들어 누군가가 'NHL' 키워드를 사용하고 이후에 'National Hockey League'를 사용하여 검색한다면, 이것은 하나의 히트로 기록함으로써 두 키워드들 간의 관련성을 끌어낼 것이다. 이 역시, 두 개의 키워드간에 형성되는 사용자의 추론능력을 포착한다. 종종 키워드는 이전의 키워드와 완전히 관련이 없을 것이니 가끔은 적절할 것이다. 만약 다음 사용자가 키워드 선택기로부터 키워드를 선택한다면, 웹 페이지에 대한 선택이 키워드와 URL과의 링크를 강화시키는 것과 동일한 방식으로 키워드 링크를 강화시킬 것이다.

다음은 책과 같은 단순한 키워드를 입력한 후 제안될 수 있는 키워드의 예시이다.

- 책 판매
- 서평
- 전문서적
- 중고서적(second hand books)
- 헌책(used books)
- 특별 발간 서적(special edition books)

모든 키워드(어구)는 정보 추구자(사용자들) 및 정보 제공자(웹 페이지 개발자)들로부터 나온다. 시간이 지남에 따라 가장 적절한 키워드가 자연스럽게 부각될 것이다.

사용자들에 의해 사용된 모든 키워드가 테이블 20의 키워드 링크 테이블(696)에 입력된다. 따라서, 사람들이 'cassettes' 대신에 'cassetes' 등의 일반적이지 않은 키워드를 입력하면 제안지는 사용자들에게 'cassettes'를 시도하도록 제안할 것이다. 따라서, 테이블 30에 'cassetes'에 대한 URL-키워드 연결의 세트를 생성할 필요가 없다. 따라서 저장공간을 절약할 수 있으며, 키워드 'cassettes'에 대한 태그된 결과 세트를 전송할 필요도 없다. 이러한 이유로 검색 엔진으로 재전송되는 데이터를 줄일 수 있다.

또 하나의 바람직한 실시예로 생각될 수 있는 것은, 테이블 30에서 유지하고 하이폴라키 키워드 연관성 및 신규 키워드 연관성을 가지는 키워드 제시지를 실행하여, 시간이 지남에 따라 키워드 연관성이 어떻게 변하는지를 시스템이 학습할 수 있다. 예를 들면, 누적 서퍼 행적에 의해 수정되어지는 테이블 30에서와 같은 방식으로, 키워드 제안지 행적은 가장 최근 키워드 링크를 저장할 수 있으며 주요 키워드 행적을 히스토리 인자에 의해 수정할 수 있다.

누적 키워드 서퍼 행적(698)은 키워드 링크 테이블(696)(테이블 10)에 키워드들 간의 링크를 강화시키기 위해 도 5의 누적 웹 페이지 서퍼 행적(170)같은 방식으로 처리된다. 만약 사용자가 매우 빨리 또 다른 키워드를 선택한다면 이전의 키워드가 키워드 서퍼 행적으로 유용하지 못하여 카운트되지 않는 것으로 되게 하기 위하여 시간변수가 포함될 수도 있다.

또한, 개개의 키워드 제안지는 각 사용자에 대하여, 그들의 개인 키워드가 링크를 저장할 수 있다. 또한, 키워드 제안지는 다수의 다양한 관심정보 타입에 기초될 수 있다. 단어 연관성은 서로 다른 문화, 국적, 직업, 연령 등의 사람들에 대해 매우 다양할 수 있다. 서로 다른 키워드 제안지는 서로 다른 그룹의 사람들에게 키워드 연관성을 포착할 수 있다. 상기한 바와 같이, 테이블 30에서 다양한 타입의 관심정보에 대해 X, Y, 및 Z 값이 기재되는 것과 같은 방식으로 테이블 10의 키워드 히트가 기재될 수 있다.

제안된 키워드의 목록을 생성하기 위해 테이블을 이용하는 방법

도 21은 제안된 키워드의 목록이 생성될 수 있는 다양한 방식을 도시한다. 일 방식은 키워드 링크 테이블(696)(테이블 10)에 X값을 순위매김에 의한 것이다. 이렇게 순위매겨진 키워드의 목록은 테이블 19의 단계

(646)에 대해 앞에서 언급한 키워드의 통상적인 검색으로부터의 키워드와 조합된다. 단계(730)로 도시된, 키워드를 제안하는 또 다른 방식은, 사용자가 입력한 키워드에 대한 인기 목록(URL의 X값)을 테이블 3의 다른 키워드의 인기 목록과 비교하는 것이다. 테이블 3의 동일한 패턴의 X값은 이러한 키워드들이 동일하다는 것을 나타낸다. 예를 들어 사용자는 'film reviews'에 대해 검색할 수 있고, 키워드 제안자는 보다 포괄적으로 검색되는 사이트의 목록을 가지는 'movie reviews'를 제안할 수 있다. 이러한 경우에 있어서, 단어 movie와 film간에 눈에 보이는 유사성은 없지만, 테이블 3에서는 이들이 공통적으로 가지는 URL 링크의 패턴의 유사성에 의해 링크된다.

키워드 제안지 목록의 유용성은 단계(744)에 의해 나타낸 바와 같이 도 5(테이블 3)의 키워드 URL 링크 테이블(172)의 키워드에 대해 상기 모든 검색이 유용한 것인지의 표시인 제안 목록상의 각 키워드를 연관시킴으로써 강화된다. 가장 많이 검색된 결과의 키워드가 부각된다.

태그된 키워드 제안 목록을 전송시키기 위한 결정

보인 테이블(168) 및 키워드 링크 테이블(696)은 웹 페이지를 태그하는 것에 관하여 상기한 것과 유사한 방식으로 어떤 키워드가 샘플에 링크될 것인지를 결정하는데 이용된다. 웹 페이지를 태그하는데 대한 결정에 있어서와 마찬가지로, 이는 그것이 반복되는 키워드(보인 테이블(168)로부터 발견)인지 여부 및 키워드 사용의 빈도(키워드 테이블(164)로부터 발견)의 빈도에 좌우될 수 있다.

다양한 콘텐츠를 판단하는 방법

상기한 바와 같이, 인터넷상에서 검색을 할 때, 다양하고 상이한 웹 페이지 목록 및 웹 페이지가 보여진다. 이러한 다양한 웹 페이지 목록의 한가지 공통된 특성은 그것들이 보여질 때 그것들이 대체로 서로 비슷한 것처럼 보인다는 것이다. 도 25에 도시된 바와 같이 다양한 목록(900)의 각각은, 비록 텍스트는 상이할 수 있을지라도 시각적으로는 동일하다. 하지만, 목록(900)보다 몇 배 더 큰 다른 목록(902)은 그래픽 콘텐츠를 포함할 수도 있어 사용자에게 보여질 때, 좀더 부각되는 것처럼 보일 수 있다. 이러한 목록은 웹 페이지 목록과 같은 콘텐츠, 또는 광고, 그림, 시설 등의 다른 콘텐츠를 포함할 수 있다.

예를 들면, 이러한 다른 콘텐츠는 키워드, 사용자 신상정보 타입(국적, 연령, 성, 직업 등) 및 일시 등에 기초하여 소정의 사용자에게 보여질 수 있다.

많은 경우에 있어서, 웹 페이지 목록과 함께 보여지는 이러한 콘텐츠는, 종래의 검색 엔진에 대하여 전송된 바와 같이 검색 시스템마다 다른 기구를 사용하여 표시 영역 내에 삽입된다. 대부분의 경우에 있어 이러한 콘텐츠들이 표시되는 기구는, 사용되는 공간에 대한 지물 등과 같은 다른 몇몇 판단기준에 기초한다. 이러한 콘텐츠를 선택하는 시스템이 작동되는 것은 하지만, 특히 콘텐츠가 자주 변경될 때에는, 어떤 콘텐츠가 보여질 것인지에 관한 정보를 끊임없이 얻어내기 어렵다. 따라서, 앞으로 논의될 본 발명의 또 다른 형태는 변경되는 콘텐츠의 행적을 쫓아 콘텐츠 제공자나 콘텐츠가 보여지게 될 시점을 동적으로 선택할 수 있도록 하는 시스템이다. 도 22에 도시된 바와 같이, 이렇게 동적으로 선택될 수 있는 콘텐츠는 단계(762)에서 뷰어에 의해 입력된 키워드 또는 신상정보 타입에 기초하여 뷰어에게 보여질 수 있다. 뒤 이은 콘텐츠 선택지 단계(764)에서, 도 25에 도시된 바와 같이 적절한 콘텐츠(902)를 웹 페이지 목록(900)과 함께 선택함에 있어 일시가 고려되며 이용된다. 웹 페이지 목록(900)을 구성하는 검색결과로 전송되는 각 콘텐츠(902)는 단계(766)에서 태그된다. 따라서, 만약 사용자가 단계(768)에서 그 콘텐츠(902)를 선택하면, 서로 연관된 콘텐츠 데이터 베이스가 이미 기술한 방식으로 서버 행적 데이터로서 업데이트될 수 있게 하기 위하여 상기 선택 결과가 정보 선택지(764)에게로 피드백된다. 그런 다음에는, 단계(770)에서, 콘텐츠(902)는 콘텐츠(900)와 일반적으로 동시에 보여진다.

이미 기술한 바와 같은 서버 행적 데이터에 더하여, 보여지는 웹 페이지 목록에 대한 사용자의 패턴에 따라, 이러한 콘텐츠가 평가하는 빈도를 결정하고, 또한, 키워드, 사용자 신상정보 등에 기초하여 콘텐츠 제공자의 콘텐츠(902)가 보여지는 방식을 변경시키기 위해, 이러한 콘텐츠 실시에도 또한 웹 페이지 개발자, 또는 콘텐츠 제공자에게 제공된다. 이러한 다이나믹한 콘텐츠 유연성을 제공하기 위하여, 변경되는 콘텐츠(902)의 행적을 추적하는데 이용되는 도 23에 기술된 세 개의 부가적인 데이터 테이블이 있다. 이 테이블들은 키워드 콘텐츠 데이터 테이블(804), 개인별 신상정보 콘텐츠 데이터 테이블(806), 및 콘텐츠 제공자 데이터 테이블(812)이다. 키워드 콘텐츠 데이터 테이블(804)은 아래의 테이블 12에 좀더 자세하게 도시되며, 그 특징은 다음과 같다.

- H는 키워드에 대한 단위기간 동안의 누적 히트수이다. 이것은 사람들이 그 키워드를 선택한 회수이다.
- N은 키워드와 관련된 소정의 콘텐츠(900)가 보여지기 위해 송출된 회수이다. 이것은 반드시 H와 같지는 않은데 이는 신상정보 타입과 연관된 콘텐츠가 키워드와 연관된 콘텐츠와는 다른 선택 인지를 가질 수 있기 때문이다. 이러한 선택 인지는 투표 또는 가격 등 매우 다양하게 변할 수 있다.
- A는 각 콘텐츠 제공자로부터의 키워드에 대한 선택 인지이다(예를 들어, 선택 인지는 상기 키워드와 연관되어 매겨진 입찰가격 \$가 될 수 있다.)
- T는 각 키워드에 대한 선택 인지의 총합이며 A의 합계이다.
- P는 각 키워드에 대해 투표, 또는 가격에 의해 결정되는 콘텐츠 가격이며, T/N 이다.(예를 들어, 이것은 그 키워드로 전송되는 콘텐츠에 대한 시간당 \$비용이 될 수 있다. - 이것은 그 키워드와 관련된 가격이다.)

키워드	1개월간의 누적 히트(H)	전송된 컨텐츠 (N)	컨텐츠 제공자1 (A1)	컨텐츠 제공자2 (A2)	합계 (T)	(P)
책						
물고기						

테이블 12 키워드 컨텐츠 데이터 세트

싱기 테이블도 역시 컨텐츠 제공자가 제공하기 위해 준비된 최대 컨텐츠값(M)을 포함할 수 있다. 소정의 키워드와 연관된 웹 페이지 목록으로 보여질 컨텐츠(902)를 구비하려는 컨텐츠 제공자의 수에는 제한이 없다.

각 국가 또는 지역에 대한 분리 테이블(12)을 구비하는 것이 가능하므로, 각각의 국가 또는 지역마다, 키워드 당 컨텐츠 값이 다를 수 있다. 게다가, 각 국가 또는 지역내에서도 서로 상이한 시간에 대해서는 상이한 컨텐츠 값이 있을 수 있다.

제공자의 컨텐츠(902)는 키워드 각각과 목표 청중을 동일시함으로써 키워드와 청중 모두를 목표로 할 수 있는 것이 가능하며, 예를 들어, 럭비라는 단어와 연관된 히트수는 럭비라는 단어에 대해 검색하는 다양한 신상정보 타입으로 나뉘어 질 수 있다. 럭비에 대한 검색의 누적수 6000이 21세이하의 520, 21세-51세는 4000, 50세이상은 520으로 조개질 수 있다. 따라서, 하나의 키워드 검색내에 있어서도 이러한 히위 클래스 각각에 대해 서로 상이한 컨텐츠 값이 있을 수 있다.

키워드 데이터 세트(804)에 부가하여, 다양한 신상정보 타입(806)에 대하여 다음과 같은 타입의 데이터 세트를 구비하는 것이 가능하다. 도 23의 키워드 컨텐츠 데이터 테이블(804)에 대하여 앞에서 언급한 키워드 대신에, 각 신상정보 타입에 대하여 같은 엔트리를 포함할 수 있다.

신상정보 타입	1개월간 누적 히트(H)	전송된 컨텐츠 (N)	컨텐츠 제공자1 (A1)	컨텐츠 제공자2 (A2)	합계 (T)	(P)
남성						
여성						
직업 등						
정의되지 않은 신상정보						

테이블 13 개인 신상정보 컨텐츠

테이블 13은 다양한 키워드에 대비되는 사람들의 특정 청중에 대한 컨텐츠(902)의 컨텐츠값을 결정하며 특정 청중을 목표로 할 수 있게 한다.

테이블 13내에 신상정보 타입의 조합, 예를 들어 남자, 전문직 또는 뉴질랜드, 여자 등과 같은 조합을 포함하는 것도 본 발명의 범위내에 포함된다. 조합된 신상정보에 대한 컨텐츠 값은 개별적 신상정보에 대한 컨텐츠 값과 다를 것이다. 컨텐츠의 값을 결정하고 컨텐츠(902)를 선택하는데 관련되는 기구는 후술되는 비와 같이 동일할 것이다.

도 23의 컨텐츠 제공자 데이터 테이블(812)은 아래에 테이블 14로서 좀 더 자세하게 도시되며, 설명, 어드레스, 광고주, 비트맵(HTML 또는 지비 애플릿 또는 그에 유사한 것) 등과 같은 컨텐츠 정보를 포함하여, 컨텐츠(902)의 각각의 다양한 아이템을 구별하기 위한 고유 번호 등을 사용할 것이다.

성명	어드레스 등	컨텐츠 정보	각 컨텐츠에 대한 고유번호
홍길동			컨텐츠 번호
			컨텐츠 번호

테이블 14

이 테이블은 컨텐츠제공자에 대한 상세사항들 예를 들면 패스워드, 지불비용에 관한 상세사항(예를 들면, 신용카드 번호 및 서명), 컨텐츠 전달(사용자에게 컨텐츠가 전송된 회수), 기타 등등을 저장한다.

상기 컨텐츠 테이블에 대한 데이터 세트는 다음과 같이 채워진다(populate). 키워드 컨텐츠 데이터 테이블(804)에 대해

- 일정 기간동안의 특정 키워드에 대한 누적 히트 수인, H를 테이블 1(800)로부터 직접 가져온다.
- N은 키워드와 관련하여 컨텐츠가 전송된 회수이다. 이것은 키워드(810)와 특정되게 연관된 사용자에게 컨텐츠(902)의 아이템이 보여질 때마다 증가된다.
- A(802)값은 각 키워드에 대한 컨텐츠제공자에 의해 선택된다. 컨텐츠제공자는 또한 그 이상을 넘으면 그 키워드에 대해 전송되도록 선택하지 않게 할 최대값(M)을 입력할 수도 있다.
- T는 키워드에 대한 총 합이며 A의 합계이다.

P는 투표 또는 가격에 의해 결정되는 각 키워드에 대한 컨텐츠값이며 T/N이다.

개인별 신상정보 컨텐츠 데이터를 증가시키는 방법

- H는 각 신상정보 타입에 대한 누적 히트수이며, 이 정보는 테이블 1로부터 직접 가져온다. (인덱스된 W의 합계)
- N은 신상정보 타입과 연관되어 전송되는 컨텐츠(902)의 아이템 개수이다. 이것은 그 신상정보 타입(810)과 관련하여 컨텐츠(902)의 아이템이 전송될 때마다 증가된다.
- A(808)값은 입력과 유사한 입력과정을 거쳐 각 신상정보 타입에 대하여 설정된다. 컨텐츠제공자는 또한 지불하거나, 투표할 수 있는 최대값(M)을 입력할 수도 있다.
- T는 각 신상정보 타입의 총합이며, A의 합계이다.

P는 각 신상정보 타입에 대한 컨텐츠 값이며 T/N이다.

컨텐츠 제공자의 상세항목 테이블을 채우는 방법

컨텐츠 제공자 상세항목(812)의 대부분은 컨텐츠 제공자들에 의해 전자적으로 입력된다. 컨텐츠 제공자의 컨텐츠(902)가 전송될 때마다, 이는 컨텐츠 제공자의 상세항목 테이블(812)에 또한 기록된다. 이것은 또한 컨텐츠(902)의 지불 또는 투표에 관하여, 클릭수(820, 822, 824, 826, 828) 및 비용을 기록할 수 있을 것이다. 이것은 추후 컨텐츠 제공자에게 전송될 전자결제 또는 테이블 작성의 기반을 형성할 것이다.

사용자에게 전송되는 컨텐츠를 선택하기 위해 데이터 세트가 이용되는 방법

도 24를 참조한 이후의 논의에 있어서, 컨텐츠(902)의 오직 단 하나의 배너만이 웹 페이지 검색 결과(900) 세트로 전송되는 것을 가정한다. 웹 페이지 결과의 각 세트로 전송되는 복수의 컨텐츠 세트가 존재한다면 같은 알고리즘이 적용된다.

키워드 및 신상정보 타입이 단계(852)에서 검색 엔진에 제공된다. 키워드 컨텐츠 데이터 테이블(804), 개인별 신상정보 컨텐츠 데이터 테이블(806)로부터의 각 컨텐츠(902)의 값이 테이블에서 P값으로부터 발견된다. 단계(862)에서 결정된 키워드 또는 신상정보 타입에 대한 가장 높은 P값은 웹 페이지 목록(900)과 함께 전송되는 컨텐츠(키워드 또는 신상정보 타입)의 타입을 결정한다. 키워드에 대한 특정값이 존재하지 않을 수도 있고 사용자기 특정 신상정보 타입을 사용하지 않을 수도 있다. 이러한 경우에는 선정되지 못한 컨텐츠 아이템의 값들이 사용될 수 있다(신상정보가 없는 사용지에 대한 테이블(13)로부터). 어떤 특정 컨텐츠 아이템(902)이 전송되는 지를 고르는 것에 대해서는 아래에 논의된다. 컨텐츠 아이템(그림, 글, 관련 프로프램 등)에 대한 상세항목은 테이블 14과 컨텐츠 제공자 상세항목 테이블(814)로부터 가져오며, 단계(868)에서 사용자에게 전송된다. 각 컨텐츠제공자에게 전송되는 컨텐츠 아이템(902)의 상세항목은 단계(870)로 나타낸 비외 같이 규칙적인 주기로 컨텐츠 제공자에게 전송된다.

키워드 또는 신상정보 컨텐츠중 어떤 것을 전송할 것인지를 결정하는 방법

전송된 컨텐츠(902)의 타입은 그것이 컨텐츠에 기초한 키워드인지 컨텐츠에 기초한 신상정보 옵션인지에 달려있다. 예를 들면, 미국 국적의 남성이 물고기에 대해 검색을 할 수 있다. 이 검색에 적용될 수 있는 값은 키워드=물고기, 신상정보=남성, 신상정보=미국, 신상정보=미국, 남성이다. 어떤 컨텐츠가 보여지게 될지를 결정할 때, 시스템은 모든 기능성에 대해 컨텐츠의 값을 비교하고 단계(862)에서 결정되는 것과 같이 가장 큰 값을 가지는 컨텐츠를 전송한다. 예를 들면, 21세 이하 남성이 '럭비'라는 키워드를 이용해 검

색을 할 수 있으며 력비와 관련된 콘텐츠에 대한 값은 뷰당 0.1임에 비하여 21세이하 남성을 목표로 한 뷰당 값은 0.2이므로, 오히려 력비 콘텐츠보다는 21세이하 남성을 목표로 한 콘텐츠가 보여지게 될 것이다. 콘텐츠 아이템(902)이 전송되는 누적 회수(N)가 목표 구역에 대한 총 누적빈도(H)와 다를 수 있다는 점이 중요하다. 이 예시에서 '력비'와 '21세이하 남성'의 검색된 회수의 누적 빈도(H)는 둘 다 1만큼씩 증기될 것이다(테이블 1을 통하여). 하지만, 콘텐츠(902)의 아이템이 보여지는 회수는 오직 '21세이하 남성' 테이블에 대해서만 증기될 것이다.(이것은 단위 뷰당 콘텐츠의 값을 결정하기 위해서 사용되는 수치이다.)

어떤 특정 콘텐츠가 전송될 지를 결정하는 방법

아래의 예시는 키워드와 관련된 콘텐츠가 어떻게 선택되는지를 나타낸다. 신상정보 타입과 관련된 콘텐츠에 대한 것과 같은 과정이다.

키워드	1개월간의 누적 히트(H)	전송된 콘텐츠의 빈도 (N)	콘텐츠 제공자1 (A1)	콘텐츠 제공자2 (A2)	합계 (T)	(P)
책	134	134		10	10	0.050
물고기	52	80	5	5	10	0.52

'책'이라는 키워드에 대해서 이 키워드가 검색될 때이다. 키워드와 관련된 유일한 정보 제공자이므로 콘텐츠 제공자(2)의 콘텐츠(902)가 보여질 것이다. 하지만, '물고기'라는 키워드에 대해서는, 콘텐츠 제공자(1 및 2)들이 그 콘텐츠를 동일한 회수로 전송할 것이다. 사용되기를 꺼하는 수준까지 올리게 될 시스템에 있어서는, 다양한 키워드 및 신상정보 타입에 대하여 입찰하는 수 많은 정보 제공자들이 존재할 수 있다.

콘텐츠의 가치를 계산하는 방법

만약 '책'이라는 키워드에 대해 예를 들어 월 \$5로 콘텐츠의 가치를 매기는 새로운 정보제공자가 있다면, 이것은 0.075까지 가치가 변할 것이고 이것은 책이라는 단어와 연관된 총합은 \$15란 것을 의미한다. 따라서, 콘텐츠 제공자(2)는 당장 시점의 66%(10/15)로 전송되어 잘 것이고 새로운 콘텐츠 제공자는 시간의 33%로 보여질 것이다. 전송되는 콘텐츠 제공자의 콘텐츠와 시간의 비는 A/T이다.

콘텐츠제공자가 데이터 테이블을 이용하는 방법

콘텐츠(902)를 입찰할 때, 콘텐츠 제공자는 테이블 (12와 13)로부터 자신의 콘텐츠를 목표로 할 키워드 또는 신상정보를 선택한다. 검색 엔진은 이전 기간(H)동안에 이 검색이 수행되었던 회수와, 그 선택에 관련되어 콘텐츠 아이템의 전송회수(N) 및 콘텐츠의 가치(P)를 자동적으로 표시한다.

새로운 콘텐츠 제공자는 선택 인자(A)를 입력하고 시스템은 새로운 총 입찰가(T)에 기초하여 새로운 가치(P)를 즉시 계산할 수 있다. 광고주는 가격($N \times (A/T)$)에 대해 믿을 수 있는 월별 뷰(view)수를 알 수 있다. 이러한 변화는 실시간으로 계산되어, 새로운 콘텐츠 제공자에게 그들의 입찰에 대해 받게될 가치 및 뷰에 그들의 입찰이 어떻게 영향을 미치게 될 것인지에 대한 표시를 제공한다. 만약 가치 및 뷰수에 광고주 동의하면 예를 들어 하루, 일주일, 한달, 등의 제한된 기간동안 그것을 입찰로 제공하는 것을 선택할 수 있다. 바람직하게는, 다른 콘텐츠 제공자의 상세항목은 공개되지 않는다. 콘텐츠 제공자는 또 그들의 콘텐츠를 팔 수 있는 최대값(M)을 입력할 수 있다. 이것은 가치가 변하더라도 너무 많은 지출을 하게 되는 것을 막는 보인으로서 콘텐츠 제공자들에게 제공된다. 가치가 너무 오르면 콘텐츠 제공자들은 그 목록에서 뺄 수가 있다(만약 P가 M보다 더 크면 A는 특정 콘텐츠 정보에 대한 입찰로 키운트되지 않는다). 가치가 떨어지면 입찰은 목록상으로 돌아오고, 따라서 안정화 기구로서 작용한다. 바람직한 실시예에 있어서, 가치한계(M)때문에 콘텐츠(902)가 목록에서 빠지면 콘텐츠 제공자는 이메일을 통해 통지받을 수 있다.

예를 들면, 도 24의 정보 제공자 상세항목 테이블(812)에 의해 나타낸 바와 같이, 콘텐츠 제공자들은 검색 엔진 소유자와 계정을 가지며 콘텐츠에 대한 계정을 주기적으로 계정 상세항목으로부터 자동적으로 계산되도록 가입하는 절차를 가진다. 각 콘텐츠 제공자에 있어, 뷰의 갯수, 뷰에 대한 비용, 클릭 횟수 및 클릭당 비용에 대한 전자적 전송은 또한 각각의 콘텐츠 제공자에게 전달되는데, 이는 상기 정보기 콘텐츠 제공자 세부항목 테이블(812)(테이블 14)에도 저장되어 있기 때문이다.

바람직한 실시예에 있어서는, 키워드 링크 테이블을 기초로하여 유사한 키워드들의 묶음('클러스터')을 식별하는 것이 가능하다. 키워드들의 클러스터를 식별하는 이유는, 콘텐츠(902)가 단지 개별적인 단어로써 보다는 단어들의 그룹으로서 목표가 될 수 있게 하기 위함이다. 키워드 '치'의 클러스터에는 '치'라는 단어와 링크되는 수 백 또는 수 천의 단어들(예를 들어, 컨버터블 자동차, 승용차, 밴)을 포함할 것이다. 키워드 집단의 크기 및 빈도를 정의하는 데에는 통계적인 클러스터 기술이 사용된다. 이것은 콘텐츠 제공자가 목표로 하는 키워드의 집단을 바탕으로 하여 편집자가 결정하는 것보다 훨씬 자동화된 과정이 되게 한다.

동일한 시스템이 키워드 집단에 대한 값을 설정하는 데 사용될 수 있다. 이러한 방법으로 단어들을 묶는 것이 관리 비용을 증기시킬 수 있기는 하지만, 이것은 계산적으로 유지할 뿐인 이니리, 키워드에 대하여

소정의 히트 수준을 넘긴 다음에이 개시된다.

오로지 콘텐츠만 검색하는 사용자는 또한 의도적으로 키워드와 관련된 콘텐츠 제공자인 검색하려고 고를 수 있다. 이 경우에 검색 결과물은 테이블 12의 A 값을 기초로할 수 있다. 가장 많이 지불하는 콘텐츠 제공자는 목록의 최상에 있을 것이다.

키워드 제시지는 또한 콘텐츠 제공자가 그가 표시하고자 하는 키워드 또는 키워드의 세트를 고르도록 도와 줄 수 있다.

검색 엔진 시스템에 대한 제어.

본 발명에 따른 검색 엔진이 웹 페이지의 순위를 매기는 방법을 비할 수 있는 많은 파라미터가 존재한다. 다음과 같은 (하기에 상술되는) 인자가 있다.

- 히스토리 인지

이것은 이전 문장에서 서술된 비와 같이 기존의 인기 목록(인기 히트 목록)의 소멸 속도를 결정짓는다. 이것은 1과 0 사이의 수다. 히스토리 인지가 크면 기존의 인기도 목록을 바꾸기 어려울 것이다. 일례로, 만일 특정 키워드에 대한 검색 속도가 급증하면, 히스토리 인지는 신생 웹 페이지를 인기도 목록에서 상층하게 할 수 있도록 작아져야 한다.

- 누적 서퍼 행적으로부터 테이블 3을 업데이트하는 빈도

이것은 사용자의 활동에 관한 정보(즉, 서퍼 행적)를 가지고 인기도 목록이 업데이트되는 빈도의 측정으로서 예를 들어, 이것은 특정한 키워드 검색의 인기도 변화의 속도에 따라 하루에 한 번 또는 한 달에 한 번 측정될 것이다.

- 표본 추출 빈도

이것은 사용자가 어떻게 검색하는지에 대한 정보를 표본 추출하는 빈도이다. 평범한 키워드인 경우에는 모든 검색을 모니터할 필요가 없다. 이것은 웹 페이지의 인기도를 정확히 결정하기 위해 모니터할 필요가 있는 모든 검색의 백분율로만 될 것이다.

- 디폴트 검색 목록의 작성(신규 웹 페이지 고공 비행 및 인기 리스트 등으로부터의 검색 결과물의 혼합)

검색 결과물이 오랜 시간 경과하였다는 것을 반영할 필요가 있으면 사용자에게 디폴트로 제공되는 웹 페이지 혼합이 변경될 수 있다.

- 콘텐츠 '히트 인지'

'콘텐츠 히트 인지'는 키워드에 대한 히트로서 기록될, 콘텐츠에 대한 히트에 주어지는 기중치를 측정하는 수단이다. 디폴트 설정은 콘텐츠에 대한 히트를 웹 페이지의 목록으로부터의 히트와 같이 동일한 수로 세는 것이다. 콘텐츠 히트의 값은 콘텐츠의 가치에 따라 전체 보다 높거나 낮게 설정될 수 있어서, 예를 들어 '콘텐츠 히트 인지'는 가치있는 키워드에 대해서는 증가될 필요가 있는데, 이는 이것이 이러한 상업적으로 가치있는 키워드를 스팸(spam)할 가능성을 감소시킬 것이기 때문이다. 콘텐츠 인지가 커질수록, 검색 결과물이 인기도 보다 가치에 더 의존하게 될 것이므로 스팸에 대한 저항도 커진다.

- 콘텐츠 입찰에 대한 기간

콘텐츠 제공자는 특정한 시간 주기 예를 들어, 한 달에 대한 소정량의 값을 입찰한다. 이러한 시간 주기는 가격의 변화율에 따라 상이해진다. 만일 가격이 급속도로 비싼다거나 매우 안정적이면, 시간 주기는 그에 부합하여 각각 단축되거나 연장될 것이다.

5 FILE="pct00046.tif"> 제출 웹 페이지당 키워드의 수

상기 수는 시스템이 신규 웹 페이지 제출을 어떻게 인식하는지에 대한 영향으로 비할 수 있다.

- 신규 웹 페이지 제출을 수용하는 동안의 시간 길이

만일 웹 페이지에 대한 제출일이 기존 웹 페이지의 제출일과 너무 인접하면 그것은 수용되지 않는다. 상기 시간의 길이는 상기 인자들 중 어느것에 따라 비할 수 있다.

- 유효 히트로 계산하는 1인(IP 어드레스 또는 사용자 ID)당 1일 검색 수

상기 숫자는 스팸의 확률을 줄이기 위해 비할 수 있다.

- 보안 테이블을 업데이트하기 전의 시간 길이

도용을 막는 보안 테이블로서, 사용자 식별의 IP 어드레스와 키워드간의 링크를 기재한다. 상기 테이블을 업데이트하는 동안의 시간 길이는 시스템을 스팸하기 어렵게 만들기 위해서 비할 수 있다.

이러한 인자의 설정은 다음에 의존하여 상이한 키워드 또는 사람들의 그룹에 따라 상이하게 될 수 있다.

- 검색이 완료되는 빈도

- 검색 빈도의 변화율

- 콘텐츠의 가치

• 콘텐츠의 가치의 변화율

상기 각 인지의 정확한 설정은 시스템이 동작을 시작하여 사용자의 행동양식에 관하여 '학습'하기 전까지는 알려지지 않을 것이다. 상이한 상황에 대한 최적의 설정은 실험에 의하여 결정될 것이다.

신용성이용가능성

기타 다른 응용

비록 본 바람직한 실시예가 컴퓨터 네트워크 상에서 인터넷을 검색하기 위해 사용가능한 소프트웨어와 관련되어 서술되었지만, 본 발명이 인간 사용자가 초기 검색 결과물의 세트로부터 결과물을 선택하는 어떠한 검색 엔진에도 용이하게 적용가능할 것이라는 것은 명백하다. 그러한 시스템은 예를 들어, LAN 또는 WAN 또는 개인 PC 상의 데이터베이스의 일부일 수도 있다.

본 발명의 가능한 다른 영역으로의 적용례가 후술된다.

인트라넷 검색 및 다른 데이터베이스 검색

현재의 인트라넷 검색은 인터넷 검색의 단점과 유사한 문제점을 가지고 있으며, 실제로 어떤 인트라넷은 특정 정보의 소스 또는 아이템이 동일하게 문제를 야기할 정도로 그 자체가 매우 큰 시스템일 수 있다. 본 발명을 그러한 적용례에 이용하는 것은 본 발명의 의도된 범위 이내의 것이다.

다른 매체 형식에 대한 검색

또한 본 발명은 사용자의 신상정보를 (페이-퍼-뷰(pay per-view), 텔레비전, 비디오, 음악등과 같은) 다른 매체 소스에 맞게 적용되도록 의도함으로써, 콘텐츠가 특정 대상자에게 전해지게 한다. 자료를 사용자에게 전달하는 데 상술한 바와 같은 검색 목록 형식(인기 목록, 고공 비행, 핫오프 기사 등)이 사용될 수 있을 것이다.

쇼핑

여기에서 서술된 검색 기술은 소비자 네트워크에서 고객을 도와 하나의 매장 또는 수많은 매장 중에서 물품을 선택할 수 있게 하는 수단이 될 수 있다. 키워드-URL 링크 테이블을 사용하는 대신에, 각 구매 요청(키워드) 후에 구매된 물품을 기록한, 구매된 키워드-물품-링크 테이블이 사용될 것이다. 또한 본 실시예는 사용자가 어느 곳에서 상품을 구입했는지를 기록한다. 매번 구매자가 물품을 구입할 때마다, 전술한 것과 동일한 기술을 사용하여 그 물품의 인기를 상승시킬 것이다.

본 실시예에서의 신상정보 타입은 사람들의 상이한 집단에 의해 형성된 구매 타입을 기록하는 데 사용될 수 있다. 예를 들어, 신상정보 타입을 선택하고 일정 범위의 사용자에게 가장 보편적인 구매 품목이 무엇인지 알 수 있으며, 자신과 상이한 신상정보 타입을 가진 사람을 위한 선물을 고르는 데 도움을 줄 것이다.

과학적 출판

본 발명으로 과학적인 데이터베이스(온라인 논문, 저널 등)를 검색하면, 무관한 정보만을 찾을지 모르는 난해하거나 심오한 분야를 검색하는데 소비되는 시간을 상당히 절약하게 될 것이다. 그러한 사용을 위한 유효 히트의 기준은 일반적으로 정보 소스의 유용성을 확립하기 위한, 상술한 연결된 시간 특성을 포함할 것이다. 본 발명을 이용하여 학술적/과학적 논문을 중재(refereeing) 및 참조하는 것은 상이한 수준 또는 사용자의 디입(예를 들어, 박사, 교수, 대학원생 등)에 따른 분류에 의해 강화될 수 있다. 이것은 사용자로 하여금 예를 들어, 특정 분야에서 저명한 기관이 어떠한 정보 소스에 관심을 갖고 있는지를 알게 해줄 수 있다. 또한 논문의 작성자로 하여금 그의 출판물이 얼마나 자주 열려되었는지 또 어떤 분야에 확장 가능한지, 그리고 그 논문이 후속 논문에서 얼마나 자주 인용되었는지 알려줄 수 있다. 사용자가 중재할 수 있는 수준의 능력을 획득하기 위해서는 여러 조직에 정식으로 등록해야 할 것이다. 또한 사용자는 자신의 검색에 대한 중재의 수준을 선택할 것이다.

온라인 도움

근래에는 세계적으로 온라인 도움 및 지원에 대한 상당한 요구가 있으며, 특히 컴퓨터/소프트웨어를 사용하는 데 있어 그러하다. 본 발명에 따르면 소프트웨어 개발자는 특정한 요청의 형태 및 빈도에 대하여 직접적으로 피드백을 얻는 한편 사용자는 종전 사용자에게 축적해 둔 이득을 얻게 됨에 의해 그러한 요구가 충족될 것이다. 상이한 신상정보 타입은 응답자가 사용자에게 대한 예를 들어, 초심자, 숙련자 등에 따라 적합한 형식으로 제공될 수 있게 할 것이다. 예를 들어, 키워드 제시지는 더욱 만족스런 답변을 이끌어낼 수 있는 키워드(질문)를 지난 검색을 제시할 것이다. 각 질문에 대한 여러가지 답변이 있을 수 있고 시스템은 그것이 최선의 답변으로 수렴될 것이라는 것을 학습하게 될 것이다.

질문 및 답변 서비스

최근의 온라인 질문/답변 프로그램은 본 발명을 통해서 작동되도록 구성될 수 있으며, 따라서 반복적으로 묻게되는 질문에 대한 답변이 종전의 질문 및 제시되는 유사한 질문의 근거가 될 수 있다.

인터넷의 다른 부분에 대한 콘텐츠 최적화

동일한 콘텐츠의 가치를 입찰하는 기구는 상기와 동일한 웹 페이지 목록뿐만이 아니라 인터넷 상의 임의의 위치에 대한 콘텐츠의 가치를 결정하는데 사용될 수 있다. 본 실시예에 있어서, 콘텐츠 제공자는 지동으로 가치를 설정하기 위한 일반 콘텐츠 공간에 대해 가치를 입찰할 것이다.

검색 엔진으로부터 신상정보 타입의 정보는 일종의 통행증(passport)으로서 사용될 수 있어, 인터넷 상의 다른 광고가 상이한 대상자에게 더 많이 전달될 수 있다. 또한 이러한 신상정보 타입의 정보는 웹 페이지

개발자가 자신의 웹 페이지를 상이한 사용자 집단을 위해 상업화하는 데 사용될 수 있다.

만남 주선 서비스

또다른 실시예에 있어서, 본 발명에 따른 시스템은 사용자 집단의 개인별 선호도(테이블 6)를 비교하여 통계적으로 분석함으로써 유사한 기호를 갖는 사람들끼리의 만남 주선 및 서로 어울리게 하는 방법으로 사용될 수 있다. 본 실시예에서 개인별 과거 선호도 테이블은 표준 상관관계 계수를 사용하여 정규화되고 서로 비교되는 것이 바람직하다. 이것을 다른 사용자와 비교하면 그들의 기호가 얼마나 유사한지에 대한 수치 지료가 제공될 것이다.

동일한 실시예가 상기 과거 선호도 테이블로부터 유사한 사람에 대한 정보를 찾는 데 사용될 수 있다. 예를 들어, 누군가가 생태 경제학에 관심이 있는 뉴질랜드에 거주하는 사람의 이름을 알려달라고 요청할 수 있고, 개인 선호도 테이블에 의해 검색될 것이다. 하지만 그러한 실시예는 일반적으로 인증된 사람이 자신의 정보를 유출하도록 허가하는 패스워드/허가 표시기를 포함할 것이며, 그러한 허가는 예를 들어, 동일하게 표시한 패스워드 및 권한의 등급을 갖는 사용자 또는 어떤 특징적인 것으로 자신을 인증한 사람으로만 제한되는, 특정 환경에서만 주어질 것이다.

본 발명이 현재에 가장 실용적이고 바람직한 실시예가 되도록 고려한 것과 연관되어 기술되었지만, 본 발명이 개시된 실시예만으로 한정되는 것은 아니다. 예를 들어, 상술된 각각의 특징은 단독으로, 또는, 다음의 청구항에 개시된 바와 같이, 서술되지는 않았지만 단독으로도 특허성이 명확한 다른 특징과 함께 조합하여 사용될 수 있다. 따라서, 본 발명은 첨부된 청구항의 사상과 범위내에 포함되는 다양한 변형예 및 등질적인 구성을 모두 포함하도록 되어있다.

(57) 청구의 범위

청구항 1

복수의 사용자 사이트를 구비한 컴퓨터 네트워크에서 서버 컴퓨터 상에 데이터베이스로서 저장된 복수의 데이터 항목 중에서 상대적인 중요성에 기중치를 부여하는 방법에 있어서,

상기 서버 컴퓨터에서 사용자 사이트로부터 키워드를 수신하는 단계;

상기 서버 컴퓨터에서, 상기 키워드에 대응하는 목록으로서, 각각의 목록은 또한 상기 데이터 항목 중의 하나와 대응하는 복수의 목록을 발생시키는 단계;

상기 복수의 목록을 상기 서버 컴퓨터로부터 상기 사용자 사이트 중의 하나에 전송하는 단계;

상기 서버 컴퓨터에서는 상기 사용자 사이트에 의해 상기 복수의 데이터 항목 중 어떤 것들이 선택되었는지를 검출하고, 상기 사용자 사이트에는, 상기 사용자 사이트에 의해 상기 대응하는 목록들이 선택되지마자 상기 데이터 항목 중 선택된 각각의 항목이 전송되는 단계; 및

상기 키워드에 관하여, 상기 데이터 항목 중에서 비선택된 항목보다 상기 데이터 항목 중에서 선택된 항목이 상대적으로 더 중요함에 가중치를 부여하여 상기 데이터베이스를 업데이트하는 단계를 포함하는 것을 특징으로 하는 방법.

청구항 2

제 1항에 있어서,

데이터를 업데이트하는 상기 단계 동안에 적용된 상기 선택된 항목의 기중치가 상기 데이터 항목의 상기 선택된 각 항목에서 소비한 시간의 양에 따라 부여되는 것을 특징으로 하는 방법.

청구항 3

제 1항에 있어서,

데이터를 업데이트하는 상기 단계 동안에 적용된 상기 선택된 항목의 기중치가 상기 데이터 항목의 상기 선택된 각 항목과 연관되어 매겨진 순위에 따라 부여되는 것을 특징으로 하는 방법.

청구항 4

제 1항에 있어서,

데이터를 업데이트하는 상기 단계 동안에 적용된 상기 선택된 항목의 기중치가 상기 데이터 항목의 상기 선택된 각 항목에서 소비한 시간의 양 및 상기 데이터 항목의 상기 선택된 각 항목과 연관되어 매겨진 순위에 따라 부여되는 것을 특징으로 하는 방법.

청구항 5

제 1항에 있어서,

상기 키워드를 수신하는 단계 이전에, 사용자를 식별하는 패스워드를 수신하는 단계; 및

상기 패스워드를 사용하여 상기 사용자기 선택된 그룹의 일원인지를 확정하는 단계를 더욱 포함하며;

상기 사용자기 상기 선택된 그룹의 일원으로 확정되는 경우에만 상기 검출하는 단계가 실행되는 것을 특징으로 하는 방법.

청구항 6

제 1항에 있어서,

상기 각 데이터 항목과 연관되는 것은 업데이트 날짜이며:

상기 복수의 목록을 발생시키는 단계는, 가장 최근에 업데이트된 상기 키워드에 대응하는 데이터 항목을 근거로 하여 복수의 목록을 발생시키는 것을 특징으로 하는 방법.

청구항 7

제 1항에 있어서,

상기 각 데이터 항목과 연관되는 것은 최근 기중치 인자 X 및 종래 기중치 인자 Y이며,

상기 복수의 목록을 발생시키는 단계는 상기 최근 기중치 인자 X 및 종래 기중치 인자 Y를 사용하여 결정된, 인기도가 가장 급상승한 데이터 항목을 근거로 하여 복수의 목록을 발생시키는 것을 특징으로 하는 방법.

청구항 8

제 1항에 있어서,

상기 키워드를 수신하는 단계 이전에 사용자를 식별하는 패스워드를 수신하는 단계를 더욱 포함하며:

상기 복수의 목록을 발생시키는 단계가 수신된 상기 키워드에 대한 상기 복수의 목록으로서, 상기 사용자와 연관되어 상기 검출하는 단계 이전에 이미 검출된, 상기 사용자와 연관된 사용자 특정 목록만을 발생시키는 것을 특징으로 하는 방법.

청구항 9

제 1항에 있어서,

상기 각 데이터 항목과 연관되는 것은 복수의 그룹이며,

상기 키워드를 수신하는 단계가 상기 그룹 중 제 1그룹의 인증도 또한 수신하고:

상기 복수의 목록을 발생시키는 단계는 상기 인증된 제 1그룹과 연관된 데이터 항목만으로부터 복수의 목록을 발생시키는 것을 특징으로 하는 방법.

청구항 10

제 9항에 있어서,

상기 수신하는 단계가 상기 그룹 중 제 2그룹의 인증을 수신하며:

상기 복수의 목록을 발생시키는 단계는 상기 인증된 제 1 및 제 2그룹의 모두와 연관된 데이터 항목만으로부터 복수의 목록을 발생시키는 것을 특징으로 하는 방법.

청구항 11

제 9항에 있어서,

상기 수신하는 단계가 상기 그룹 중 제 2그룹의 인증을 수신하며:

상기 복수의 목록을 발생시키는 단계는 상기 인증된 제 1 또는 제 2그룹의 어느 하나와 연관된 데이터 항목만으로부터 복수의 목록을 발생시키는 것을 특징으로 하는 방법.

청구항 12

제 10항에 있어서,

상기 수신하는 단계가 상기 그룹 중 제 3그룹의 인증을 수신하며:

상기 복수의 목록을 발생시키는 단계는 상기 인증된 제 1, 제 2 및 제 3그룹의 모두와 연관된 데이터 항목만으로부터 복수의 목록을 발생시키는 것을 특징으로 하는 방법.

청구항 13

제 11항에 있어서,

상기 수신하는 단계가 상기 그룹 중 제 3그룹의 인증을 수신하며:

상기 복수의 목록을 발생시키는 단계가 상기 제 1, 제 2 또는 제 3그룹 중의 어느 하나와 연관된 데이터 항목만으로부터 복수의 목록을 발생시키는 것을 특징으로 하는 방법.

청구항 14

제 1항에 있어서,

상기 키워드가 허가된 키워드인지를 결정하는 단계를 더욱 포함하며:

상기 복수의 목록을 발생시키는 단계는 상기 키워드가 허가된 키워드인 경우에만 실행되는 것을 특징으로 하는 방법.

청구항 15

제 1항에 있어서,

상기 검출하는 단계가 소정 시간 간격동안에 각 사용자 사이트로부터 한 번씩만 각 키워드를 검출하는 것을 특징으로 하는 방법.

청구항 16

제 15항에 있어서,

상기 각 키워드를 한 번 검출하는 단계가,

각 사용자와 식별자를 연관시키는 단계; 및

상기 식별자를 이용하여, 상기 복수의 각 사용자 사이트로부터 이미 입력된 키워드를 추적하는 단계를 포함하는 것을 특징으로 하는 방법.

청구항 17

제 15항에 있어서,

상기 복수의 목록을 발생시키는 단계는 상기 복수의 목록을 결정하는 데 각 키워드와 연관된 히스토리 인지를 사용하는 것을 특징으로 하는 방법.

청구항 18

제 16항에 있어서,

상기 히스토리 인지는 0 이상 1 이하인 것을 특징으로 하는 방법.

청구항 19

제 1항에 있어서,

상기 각 데이터 항목과 연관되는 것은 업데이트 날짜이며;

상기 복수의 목록을 발생시키는 단계는 가장 최근에 업데이트된 상기 키워드에 대응하는 데이터 항목 및 가장 인기있는 데이터 항목을 근거로 하는 복수의 목록을 발생시키는 것을 특징으로 하는 방법.

청구항 20

제 1항에 있어서,

상기 각 데이터 항목과 연관되는 것은 최근 기중치 인자 X 및 종래 기중치 인자 Y이며;

상기 복수의 목록을 발생시키는 단계는, 상기 최근 기중치 인자 X 및 종래 기중치 인자 Y를 사용하여 결정된, 인기도가 가장 급상승한 데이터 항목 및 가장 인기 있는 데이터 항목을 근거로 하여 복수의 목록을 발생시키는 것을 특징으로 하는 방법.

청구항 21

제 1항에 있어서,

상기 키워드를 수신하는 단계 이전에 사용자를 인증하는 패스워드를 수신하는 단계를 더욱 포함하며;

상기 복수의 목록을 발생시키는 단계는, 수신된 상기 키워드에 대한 상기 복수의 목록으로서, 상기 사용자와 연관된 상기 검출하는 이전의 단계에서 이미 검출된 상기 사용자와 연관된 사용자 특정 목록 및 가장 인기있는 데이터 항목의 조합을 발생시키는 것을 특징으로 하는 방법.

청구항 22

제 1항에 있어서,

각 데이터 항목과 연관되는 것은 최근 기중치 인자 X 및 종래 기중치 인자 Y이며;

상기 복수의 목록을 발생시키는 단계는, 상기 최근 기중치 인자 X 및 종래 기중치 인자 Y를 사용하여 결정된, 인기도가 가장 급상승한 데이터 항목 및 가장 최근에 업데이트된 상기 키워드에 대응하는 데이터 항목을 근거로 하여 복수의 목록을 발생시키는 것을 특징으로 하는 방법.

청구항 23

제 1항에 있어서,

상기 키워드를 수신하는 단계 이전에 사용자를 식별하는 패스워드를 수신하는 단계를 더욱 포함하며;

상기 복수의 목록을 발생시키는 단계는, 수신된 상기 키워드에 대한 상기 복수의 목록으로서, 가장 최근에 업데이트된 상기 키워드에 대응하는 데이터 항목을 근거로 하여, 상기 사용자와 연관된 상기 검출하는 이전의 단계에서 이미 검출된, 상기 사용자와 연관된 상기 사용자 특정 목록의 조합을 발생시키는 것을 특징으로 하는 방법.

청구항 24

제 1항에 있어서,

상기 키워드를 수신하는 단계 이전에 사용자를 식별하는 패스워드를 수신하는 단계를 더욱 포함하며;

상기 복수의 목록을 발생시키는 단계가 수신된 상기 키워드에 대한 상기 복수의 목록으로서, 상기 최근 기

중치 인자 X 및 싱기 종래 기중치 인자 Y를 사용하여 결정되는 인기도가 가장 급상승한 데이터 항목을 근거로 하여, 사용자와 연관된 싱기 검출하는 단계 이전에 이미 검출된, 싱기 사용자와 연관된 싱기 사용자 특정 목록의 조합을 발생시키는 것을 특징으로 하는 방법.

청구항 25

제 1항에 있어서,

싱기 각 데이터 항목과 연관되는 것은 업데이트 날짜이며:

싱기 복수의 목록을 발생시키는 단계는, 가장 최근에 업데이트된 싱기 키워드에 대응하는 데이터 항목만을 근거로 하여 싱기 복수의 목록을 발생시키는 것을 특징으로 하는 방법.

청구항 26

제 1항에 있어서,

싱기 각 데이터 항목과 연관되는 것은 최근 기중치 인자 X 및 종래 기중치 인자 Y이며:

싱기 복수의 목록을 발생시키는 단계는, 싱기 최근 기중치 인자 X 및 싱기 종래 기중치 인자 Y를 사용하여 결정되는 인기도가 가장 급상승한 데이터 항목만을 근거로 하여 싱기 복수의 목록을 발생시키는 것을 특징으로 하는 방법.

청구항 27

제 1항에 있어서,

싱기 각 데이터 항목과 연관되는 것은 업데이트 날짜이며:

싱기 복수의 목록을 발생시키는 단계는, 가장 인기있는 데이터 항목만을 근거로 하여 싱기 복수의 목록을 발생시키는 것을 특징으로 하는 방법.

청구항 28

복수의 사용자 사이트 및 개발자 사이트를 포함하는 컴퓨터 네트워크에서, 서버 컴퓨터 상에 데이터베이스를 올리는 방법에 있어서,

싱기 개발자 사이트로부터 싱기 데이터베이스 내에 복수의 데이터 항목을 입력하는 단계로서, 싱기 데이터 항목은 복수의 연관된 키워드의 연관된 식별자를 포함한 싱기 데이터베이스 내에 입력되는 싱기 입력 단계; 및

복수의 사용자 행적을 입력하여 싱기 데이터베이스를 업데이트하는 단계로서, 싱기 각 사용자 행적이 싱기 데이터 항목과 싱기 연관된 키워드 중의 하나를 식별하여 싱기 각 사용자 행적은 연관된 키워드에 관하여 연관된 데이터 항목의 상대적인 중요성을 증가시키도록 하는 싱기 업데이트 단계를 포함하는 것을 특징으로 하는 방법.

청구항 29

제 28항에 있어서,

싱기 복수의 데이터 항목을 입력하는 단계는, 생성된 날짜와 업데이트 날짜 중의 하나를 싱기 연관된 식별자로서 포함하는 것을 특징으로 하는 방법.

청구항 30

제 28항에 있어서,

싱기 복수의 데이터 항목을 입력하는 단계는, 개발자 사이트 식별자를 싱기 연관된 식별자로서 포함하는 것을 특징으로 하는 방법.

청구항 31

제 30항에 있어서,

싱기 개발자 사이트 식별자는 싱기 개발자 사이트가 싱기 개발자 사이트에 의해 입력된 싱기 데이터 항목에 대하여 싱기 업데이트하는 싱기 단계 동안에 사용되는 것을 방지하는 데 사용되는 것을 특징으로 하는 방법.

청구항 32

제 28항에 있어서,

싱기 각 사용자 행적은 사용자 테이블을 업데이트하는 데 사용되는 사용자 식별자를 포함하는 것을 특징으로 하는 방법.

청구항 33

제 17항에 있어서,

싱기 사용자 테이블은 사용자에게 의해 실행되는 복수의 상이한 검색 타입에 대응하는 사용자 행적을 포함하는 것을 특징으로 하는 방법.

청구항 34

제 33항에 있어서,

사용자에 의해 수행된 상기 검색 타입 중의 하나는 가장 최근에 업데이트된 키워드 검색에 대응하는 것을 특징으로 하는 방법.

청구항 35

제 17항에 있어서,

특정 사용자는 복수의 사용자 식별지(각 사용자 식별지는 상이한 사용자 신상정보에 대응)를 구비하는 것을 특징으로 하는 방법.

청구항 36

제 28항에 있어서,

상기 업데이트하는 단계는 복수의 특성을 포함한 신상정보를 갖는 사용자에 의해 실행되며, 상기 업데이트 동작은, 각 복수의 특성에 관하여 상기 데이터베이스를 업데이트하는 것을 특징으로 하는 방법.

청구항 37

제 28항에 있어서,

각 데이터 항목과 연관되는 것은 복수의 그룹 및, 상기 복수의 그룹 중 어떤 그룹과 연관된 각각의 사용자이며;

상기 복수의 그룹 중 소정 그룹 내의 일 사용자에 대한 상기 업데이트하는 단계는 상기 복수의 그룹 중 상기 소정 그룹에 관한 데이터베이스를 업데이트하게 되는 것을 특징으로 하는 방법.

청구항 38

제 28항에 있어서,

각 데이터 항목과 연관되는 것은 복수의 그룹 및, 상기 복수의 그룹 중 어떤 그룹과 연관된 각 사용자이며;

상기 복수의 그룹 중 소정 그룹 내의 일 사용자에 대한 상기 업데이트하는 단계는 상기 복수의 그룹 중 단지 상기 소정 그룹에 관하여만 데이터베이스를 업데이트하게 되는 것을 특징으로 하는 방법.

청구항 39

제 28항에 있어서,

상기 사용자 행적은, 사용자 사이트로 전송되어 선택된 데이터 항목에서 사용자에 의해 소비된 시간의 양에 따라 발생될 것을 특징으로 하는 방법.

청구항 40

제 39항에 있어서,

상기 사용자 행적은 상기 사용자가 상기 선택된 데이터 항목에서 소정 시간 동안용 소비한 경우에, 발생될 것을 특징으로 하는 방법.

청구항 41

제 28항에 있어서,

상기 사용자 행적은 사용자에게 전송되어 선택된 데이터 항목과 연관되어 매겨진 순위에 따라 발생될 것을 특징으로 하는 방법.

청구항 42

제 28항에 있어서,

상기 사용자 행적은 사용자 사이트로 전송되어 선택된 데이터 항목에서 사용자에 의해 소비된 시간의 양 및 선택된 데이터 항목과 연관되어 매겨진 순위에 따라 발생될 것을 특징으로 하는 방법.

청구항 43

제 28항에 있어서,

상기 업데이트하는 단계 이전에 사용자를 식별하는 패스워드를 수신하는 단계; 및

상기 패스워드를 사용하여 사용자가 선택된 그룹의 일원인지를 결정하는 단계를 더욱 포함하며;

상기 사용자가 상기 선택된 그룹의 일원이 아닌 경우에는 상기 업데이트하는 단계는, 발생되지 않을 것을 특징으로 하는 방법.

청구항 44

제 43항에 있어서,

상기 사용자 행적은 상기 사용자가 상기 선택된 그룹의 일원이며 상기 선택된 데이터 항목에서 소정의 시간을 소비한 경우에 발생될 것을 특징으로 하는 방법.

청구항 45

제 28항에 있어서,

상기 업데이트하는 단계는 소정 시간 간격동안에 각 사용지에 대하여 각 키워드만을 한 번 업데이트하는 것을 특징으로 하는 방법.

청구항 46

제 45항에 있어서,

상기 각 키워드를 한 번 업데이트하는 단계는,

각 사용자와 식별자를 연관시키는 단계; 및

상기 식별자를 사용하여, 각 사용지에 의해 입력된 키워드를 추적하는 단계를 포함하는 것을 특징으로 하는 방법.

청구항 47

복수의 개별지 사이트를 구비한 컴퓨터 네트워크에서, 서버 컴퓨터로부터 사용자 사이트로 전송되는 다른 정보와 함께 제공할 콘텐츠를 결정하는 방법에 있어서,

상기 복수의 각 개별지 사이트로부터 콘텐츠, 개별지 식별자, 키워드 및 키워드 선택 인자를 포함하여 이루어지는 각각의 콘텐츠 목록을 획득하는 단계;

획득된 상기 키워드들 중에서 상이한 콘텐츠 목록에 대하여 동일한 특정 키워드를 결정하는 단계; 및

상기 콘텐츠 목록의 다른 목록과 연관된 다른 콘텐츠보다는, 상기 콘텐츠 목록 중의 하나와 연관된 특정 콘텐츠를 전송할지 여부를 결정하는 데 상기 키워드 선택 인자를 사용하는 것을 특징으로 하는 방법.

청구항 48

제 47항에 있어서,

각 콘텐츠 목록에 대한 신상정보 및 신상정보 선택 인자를 획득하는 단계; 및

획득된 상기 신상정보로부터 상이한 콘텐츠 목록에 대하여 동일한 특정 신상정보를 결정하는 단계를 더욱 포함하며;

상기 콘텐츠 목록의 다른 목록과 연관된 다른 콘텐츠보다는, 상기 콘텐츠 목록 중의 하나와 연관된 특정 콘텐츠를 전송할지 여부를 결정하는 데 상기 신상정보 선택 인자를 사용하는 것을 특징으로 하는 방법.

청구항 49

복수의 사용자에게 복수의 콘텐츠 제공자의 콘텐츠를 전자적으로 표시하기 위한 백분율링을 결정하는 방법에 있어서,

상기 복수의 각 콘텐츠 제공자로부터, 연관된 특정 키워드와, 콘텐츠의 일부를 표시하기 위해 상기 콘텐츠 제공자가 상기 특정 키워드에 연관시키는 가치에 대응하는 입찰량을 수신하는 단계;

콘텐츠의 각 부분이 상기 복수의 사용자에게 전자적으로 표시될 백분율링을 결정하기 위해, 동일한 키워드에 대한 입찰량을 상호연관시키는 단계; 및

상기 백분율링을 부분적으로 근거로하여, 상기 콘텐츠의 각 부분을 표시하는 단계를 포함하는 것을 특징으로 하는 방법.

청구항 50

제 49항에 있어서,

상기 백분율링은 상기 특정 키워드가 수신되는 동인의 기간을 근거로 하는 것을 특징으로 하는 방법.

청구항 51

제 49항에 있어서,

상기 백분율링은 상기 특정 키워드의 전송 회수를 근거로 하는 것을 특징으로 하는 방법.

청구항 52

제 49항에 있어서,

상기 특정 키워드에 대한 부가적인 입찰량을 수신하는 단계;

각 콘텐츠 개별지의 콘텐츠가 상기 복수의 사용자에게 전자화되어 표시되는 업데이트된 백분율링을 결정하기 위해 입찰량 및 부가적인 입찰량을 다시 상호연관시키는 단계; 및

업데이트된 백분율링을 부분적으로 근거로 하여, 상기 콘텐츠의 각 부분을 표시하는 단계를 더욱 포함하는 것을 특징으로 하는 방법.

청구항 53

제 52항에 있어서,

상기 업데이트된 백분율링은 상기 특정 키워드가 수신되는 동안의 기간을 근거로 하는 것을 특징으로 하는 방법.

청구항 54

제 52항에 있어서,

상기 업데이트된 백분율링은 상기 특정 키워드의 전송 회수를 근거로 하는 것을 특징으로 하는 방법.

청구항 55

제 49항에 있어서,

상기 입찰량 중의 하나를 대신하여 특정 키워드에 대한 변경된 입찰량을 수신하는 단계:

각 콘텐츠 개발자의 콘텐츠가 상기 복수의 사용자에게 전자적으로 표시되는 업데이트된 백분율링을 결정하기 위해 상기 입찰량 및 변경된 입찰량을 다시 상호연관시키는 단계: 및

업데이트된 백분율링을 부분적으로 근거로 하여, 상기 콘텐츠의 각 부분을 표시하는 단계를 더욱 포함하는 것을 특징으로 하는 방법.

청구항 56

제 55항에 있어서,

상기 업데이트된 백분율링은 상기 특정 키워드가 수신되는 동안의 기간을 근거로 하는 것을 특징으로 하는 방법.

청구항 57

제 55항에 있어서,

상기 업데이트된 백분율링은 상기 특정 키워드의 전송 회수를 근거로 하는 것을 특징으로 하는 방법.

청구항 58

복수의 사용자에게 복수의 콘텐츠 제공자의 콘텐츠를 전자적으로 표시하기 위한 백분율링을 결정하는 방법에 있어서,

상기 복수의 각 콘텐츠 제공자 각각으로부터, 연관된 신상정보 및, 상기 각 콘텐츠 제공자가 콘텐츠 일부의 표시를 위해 특정 키워드와 연관시킨 가치에 대응하는 입찰량을 수신하는 단계:

상기 콘텐츠의 각 부분이 상기 복수의 사용자에게 전자적으로 표시될 백분율링을 결정하기 위하여 동일한 신상정보에 대한 모든 입찰량을 상호연관시키는 단계: 및

백분율링을 부분적으로 근거로하여 상기 콘텐츠의 각 부분을 표시하는 단계를 포함하는 것을 특징으로 하는 방법.

청구항 59

제 58항에 있어서,

상기 백분율링은 상기 신상정보가 수신되는 동안의 기간을 근거로 하는 것을 특징으로 하는 방법.

청구항 60

제 58항에 있어서,

상기 백분율링은 상기 신원정보의 전송 회수를 근거로 하는 것을 특징으로 하는 방법.

청구항 61

제 58항에 있어서,

상기 신원정보에 대한 부가적인 입찰량을 수신하는 단계:

각 콘텐츠 개발자의 콘텐츠가 상기 복수의 사용자에게 전자적으로 표시되는 업데이트된 백분율링을 결정하기 위해 상기 입찰량 및 상기 부가적인 입찰량을 다시 상호연관시키는 단계: 및

상기 업데이트된 백분율링을 부분적으로 근거로하여 상기 콘텐츠의 각 부분을 표시하는 단계를 더욱 포함하는 것을 특징으로 하는 방법.

청구항 62

제 61항에 있어서,

상기 업데이트된 백분율링은 상기 신상정보가 수신되는 동안의 기간을 근거로 하는 것을 특징으로 하는 방법.

청구항 63

제 61항에 있어서,

상기 업데이트된 백분율링은 상기 신원정보의 전송 회수를 근거로 하는 것을 특징으로 하는 방법.

청구항 64

컨텐츠 제공자가 제공하는 복수의 상이한 컨텐츠 중에서 어느 컨텐츠를 전자적으로 일 사용자에게 표시할 지 여부를 결정하는 방법에 있어서,

상기 복수의 각 컨텐츠 제공지로부터 상기 각 컨텐츠 제공지가 컨텐츠의 각 부분을 표시하기 위해 적어도 하나의 키워드 또는 사용자 그룹의 신상정보와 연관시킨 기치에 대응하는 입찰량을 수신하는 단계;

어떤 컨텐츠가 상이한 각 키워드 및 신상정보에 대한 가장 큰 백분율량을 갖는 지를 결정할 수 있도록 하기 위해서, 서로 상이한 각 입찰량에 대한 상이한 키워드 및 신상정보에 모든 입찰량을 상호연관시키는 단계;

상기 사용자 그룹의 신상정보 내에 포함되는 일 사용자로부터 일 키워드를 수신하는 단계; 및

일 키워드 및 사용자의 신상정보를 근거로하여 가장 높은 대응 입찰량을 갖는 컨텐츠의 각 부분을 일 사용자에게 전송하는 단계를 포함하는 것을 특징으로 하는 방법.

청구항 65

복수의 사용자 사이트를 구비한 컴퓨터 네트워크에서 서버 컴퓨터 상의 데이터베이스 내에 저장된 복수의 키워드의 상대적 중요성에 가중치를 부여하는 방법에 있어서,

상기 서버 컴퓨터에서 사용자 사이트로부터 초기 키워드를 수신하는 단계;

상기 서버 컴퓨터에서 상기 초기 키워드에 대응하는 복수의 관련 키워드를 발생시키는 단계;

상기 서버 컴퓨터로부터 상기 사용자 사이트의 하이에 상기 복수의 관련 키워드를 전송하는 단계;

상기 서버 컴퓨터에서 상기 복수의 관련 키워드 중 어느 키워드가 상기 사용자 사이트에 의해 선택되는지를 검출하는 단계; 및

상기 선택되지 않은 키워드와 초기 키워드간의 관련성보다 상기 선택된 키워드와 상기 초기 키워드간의 관련성에 더 큰 가중치를 주도록 상기 데이터베이스를 업데이트하는 단계를 포함하는 것을 특징으로 하는 방법.

청구항 66

제 65항에 있어서,

상기 복수의 사이트의 상기 사용자를 한 데 묶은 복수의 신상정보가 존재하며, 상기 업데이트하는 단계가 상기 선택된 키워드에 대한 상기 복수의 신상정보 중의 상이한 신상정보에 대하여 상이한 가중치를 부여하는 것을 특징으로 하는 방법.

청구항 67

제 65항에 있어서,

상기 복수의 관련 키워드를 발생시키는 단계가 상기 복수의 관련 키워드를 발생시킬 때, 초기 키워드를 제출한 사용자 사이트와 연관된 신상정보도 고려하는 것을 특징으로 하는 방법.

청구항 68

제 65항에 있어서,

상기 복수의 관련 키워드는 자주 사용되는 키워드일수록 더 높은 계층에 있는 것과 같은 방식으로 계층 디렉토리로서 저장되는 것을 특징으로 하는 방법.

청구항 69

제 65항에 있어서,

부분적으로 상기 복수의 관련 키워드는 상기 초기 키워드의 잘못된 철자로부터 발생하는 것을 특징으로 하는 방법.

청구항 70

제 65항에 있어서,

부분적으로 상기 복수의 관련 키워드는 상기 초기 키워드를 포함하는 어구로부터 발생하는 것을 특징으로 하는 방법.

청구항 71

제 65항에 있어서,

부분적으로 상기 복수의 관련 키워드는 특정 사용자로부터의 과거 선호도 정보를 사용하여 발생하는 것을 특징으로 하는 방법.

청구항 72

제 65항에 있어서,

부분적으로 상기 복수의 관련 키워드는 가장 최근에 업데이트된 키워드 및 가장 인기있었던 키워드의 목록

으로부터 발생되는 것을 특징으로 하는 방법.

청구항 73

제 65항에 있어서,

각 키워드와 연관되는 것은 복수의 그룹이며;

상기 키워드를 수신하는 상기 단계는 상기 그룹 중 제 1그룹의 인증도 또한 수신하며;

상기 복수의 관련 키워드를 발생시키는 단계는 상기 인증된 제 1그룹과 연관된 키워드만으로부터 상기 복수의 관련 키워드를 발생시키는 것을 특징으로 하는 방법.

청구항 74

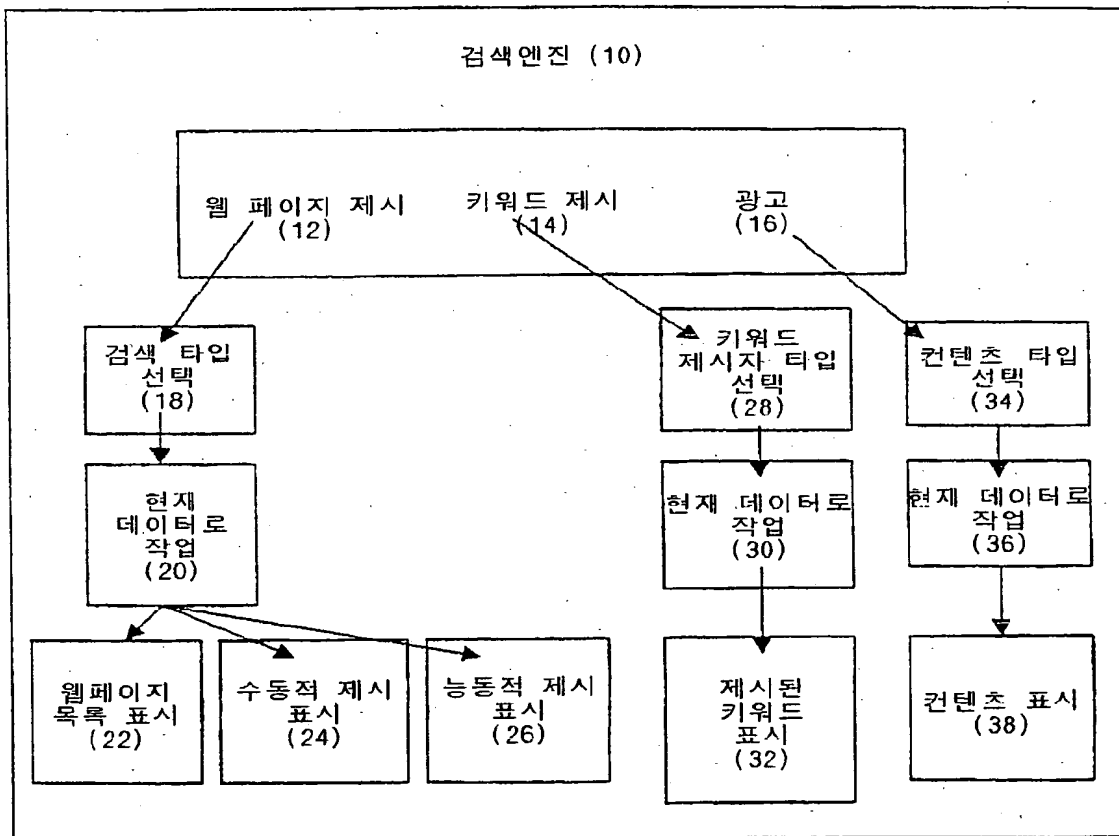
제 73항에 있어서,

상기 키워드를 수신하는 단계는 상기 그룹 중 제 2그룹의 인증을 수신하며;

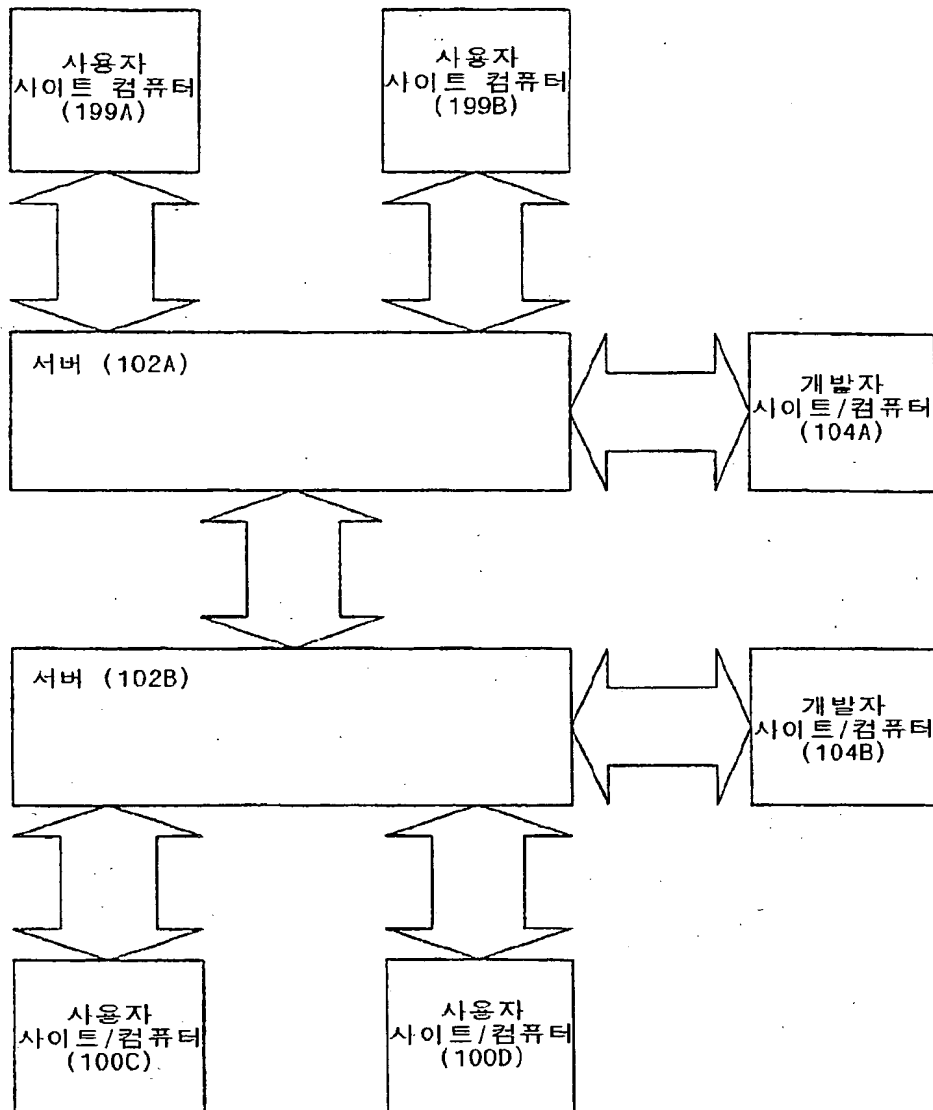
상기 복수의 관련 키워드를 발생시키는 단계는 상기 인증된 제 1 및 제 2그룹의 모두와 연관된 키워드만으로부터 상기 복수의 관련 키워드를 발생시키는 것을 특징으로 하는 방법.

도면

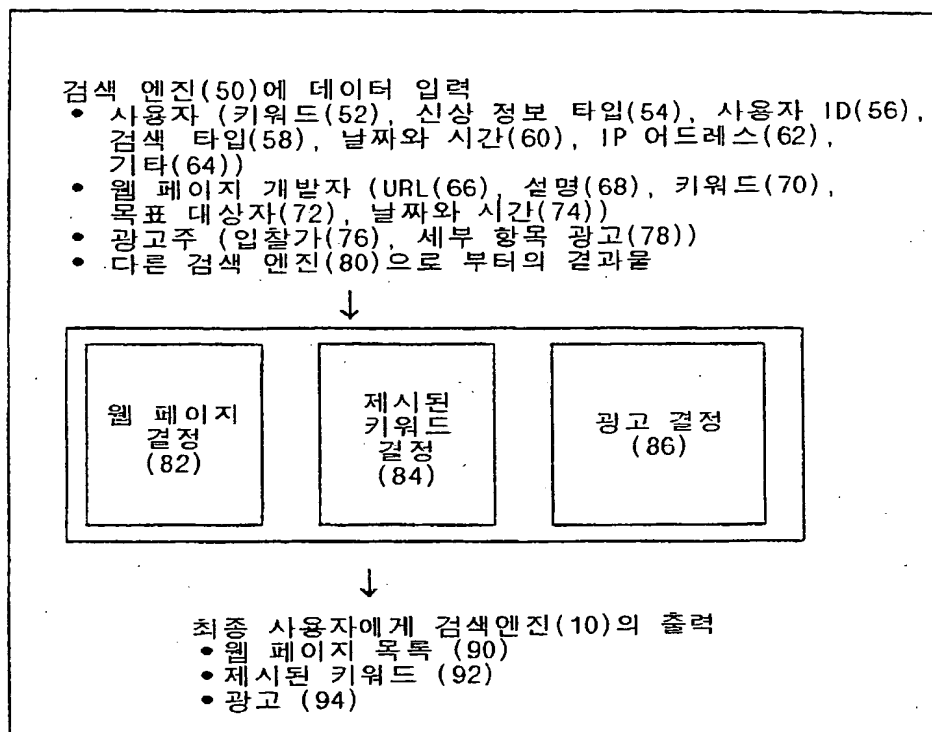
도면 1a



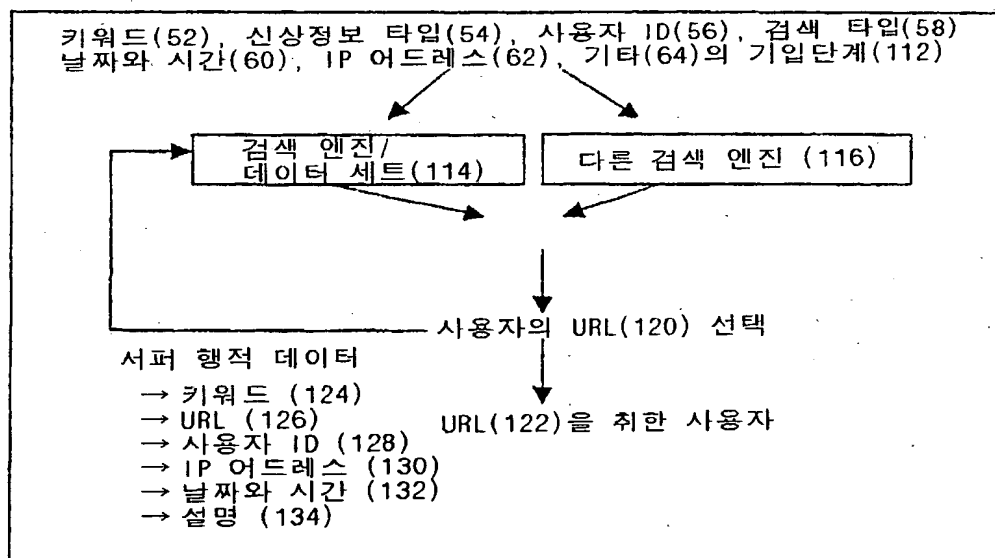
도면 1b



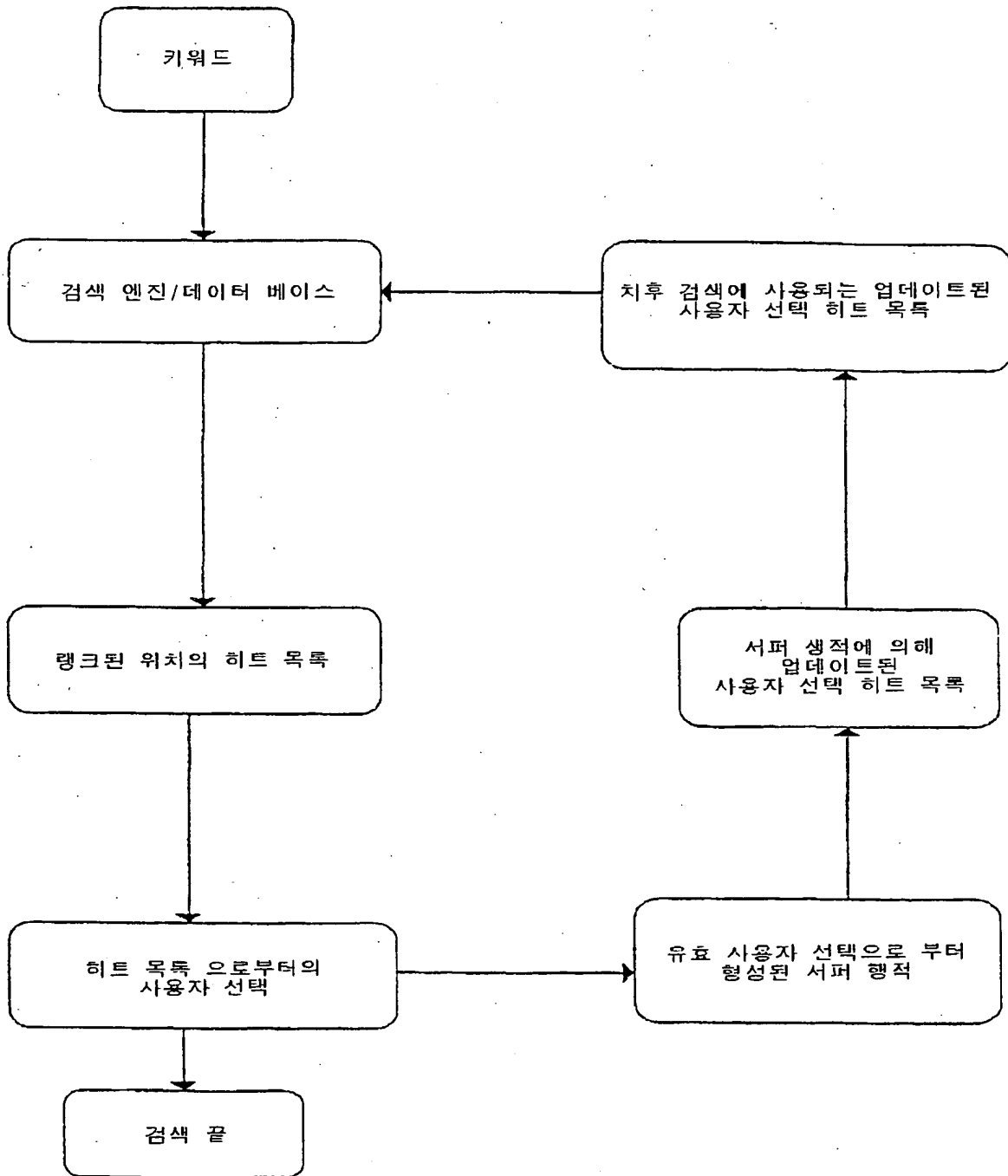
도면2



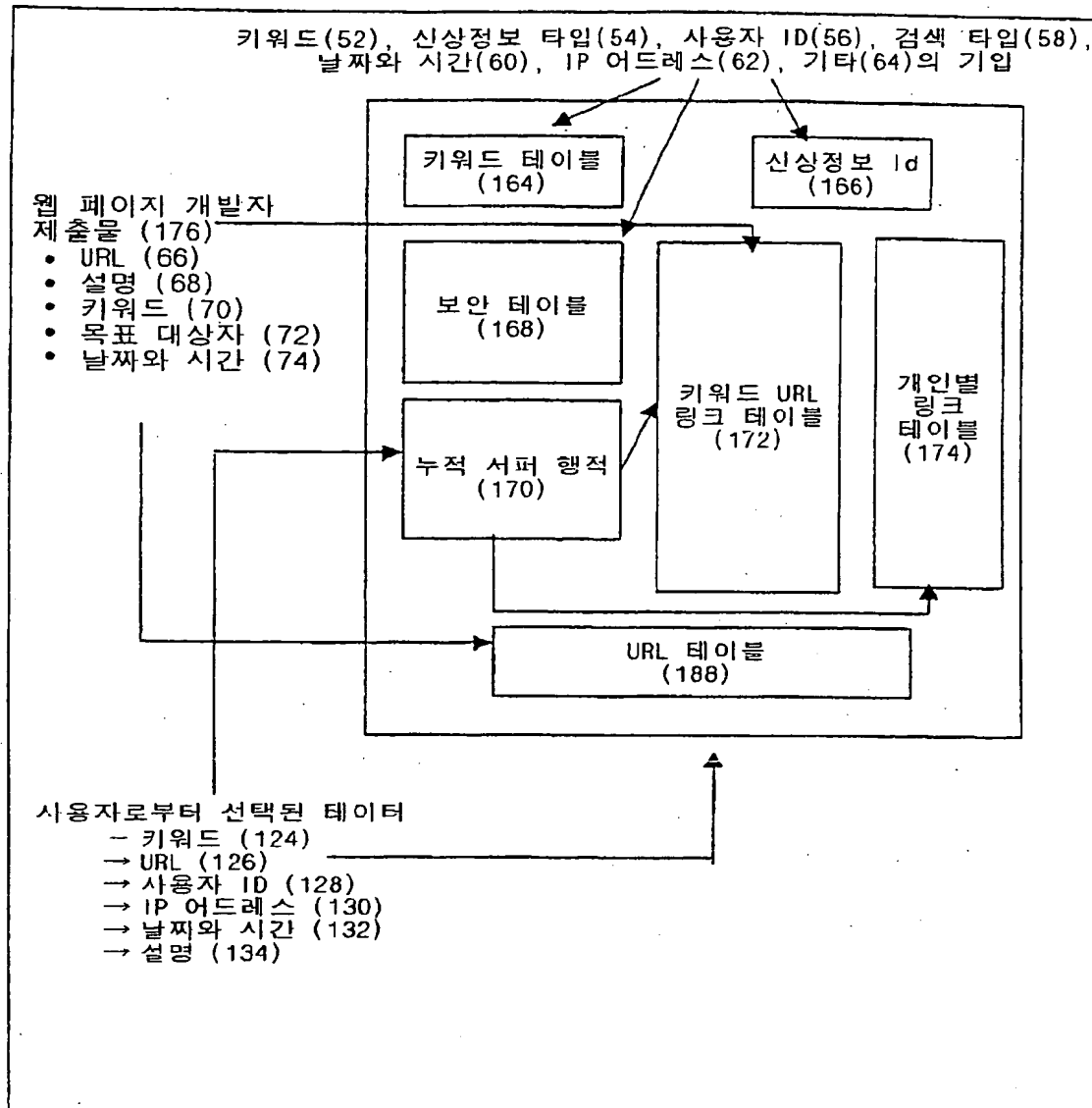
도면3a



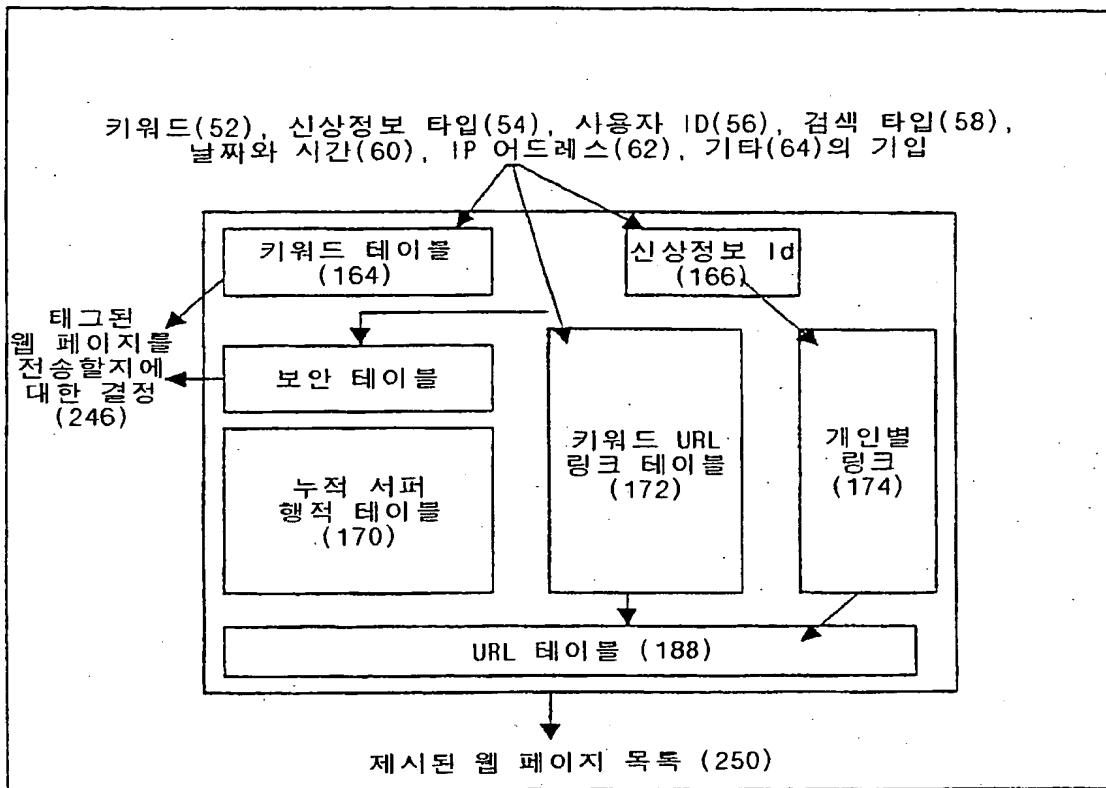
도면3b



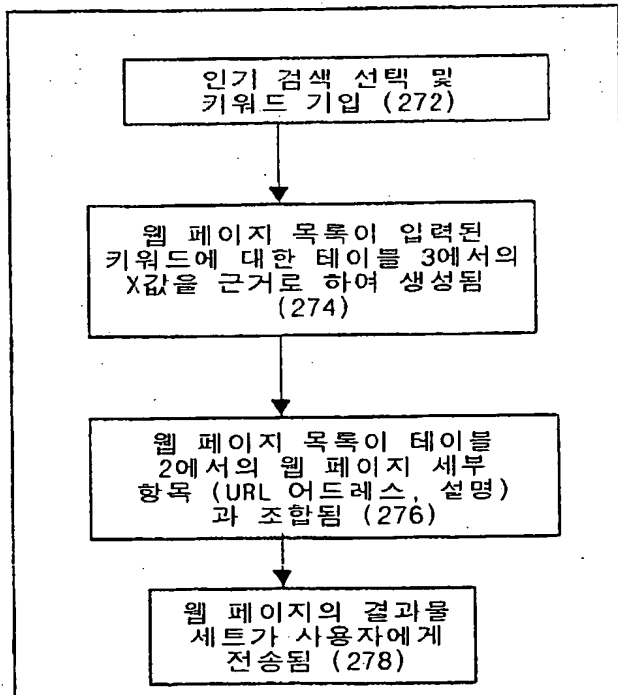
도면4



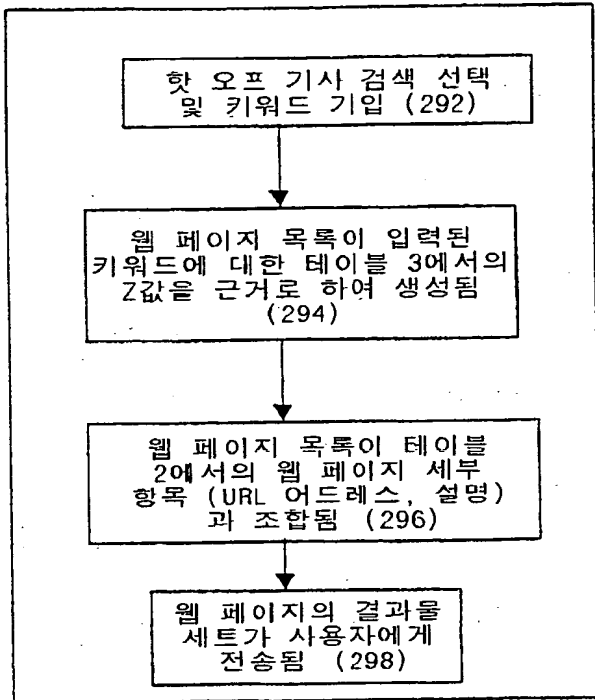
도면5



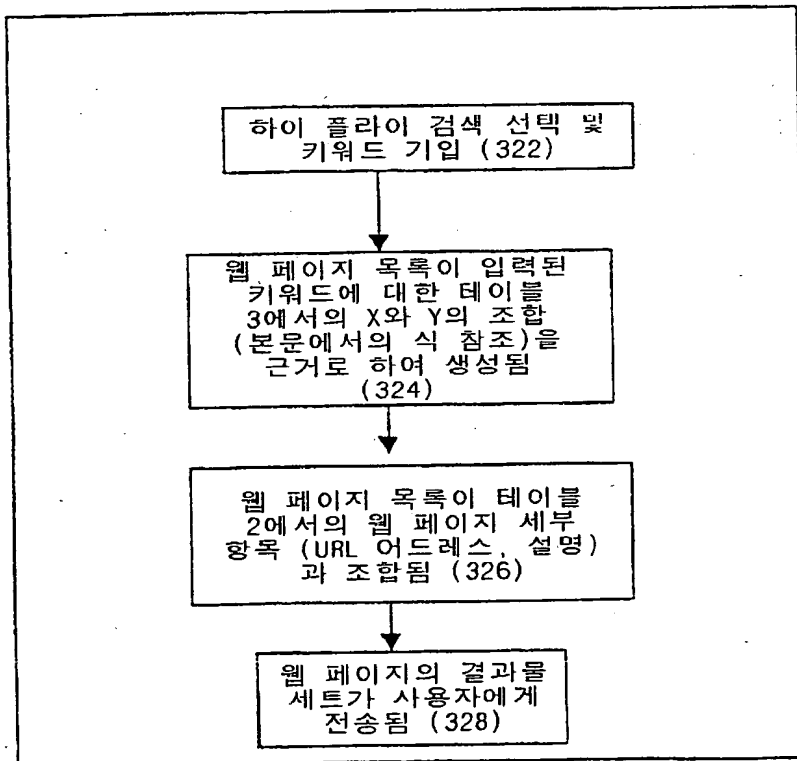
도면6



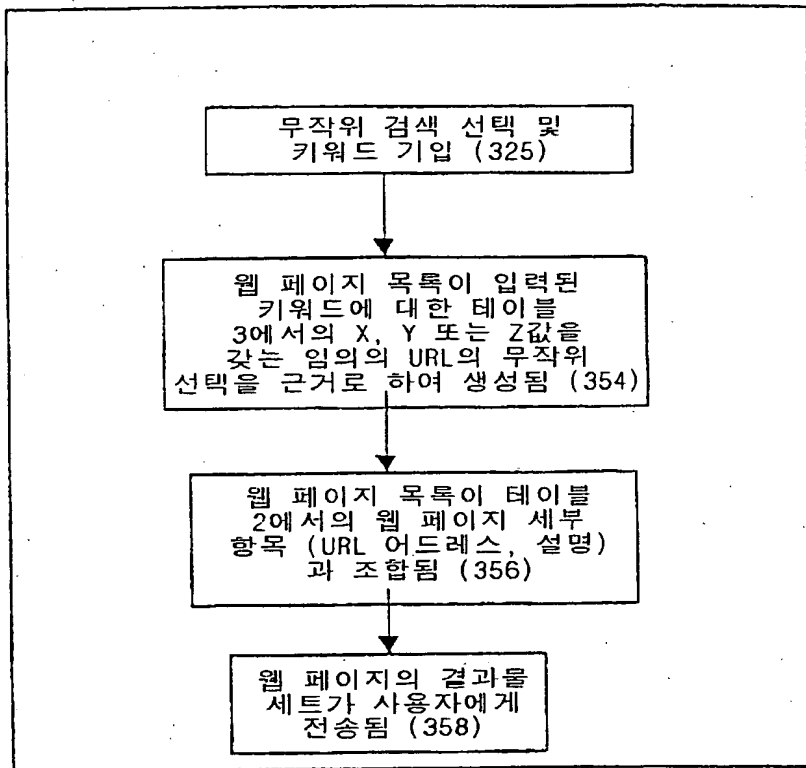
도면7



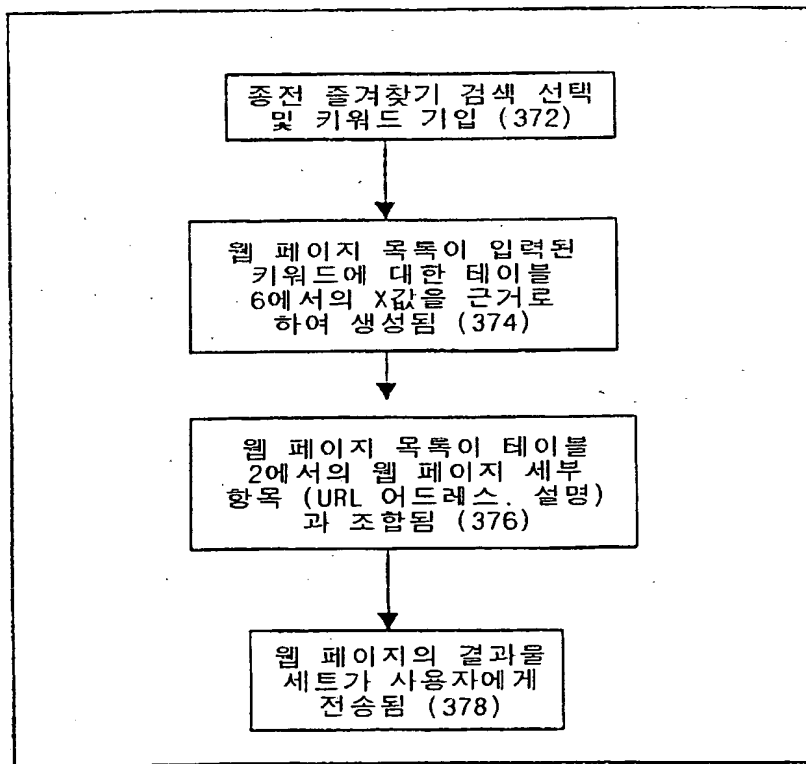
도면8



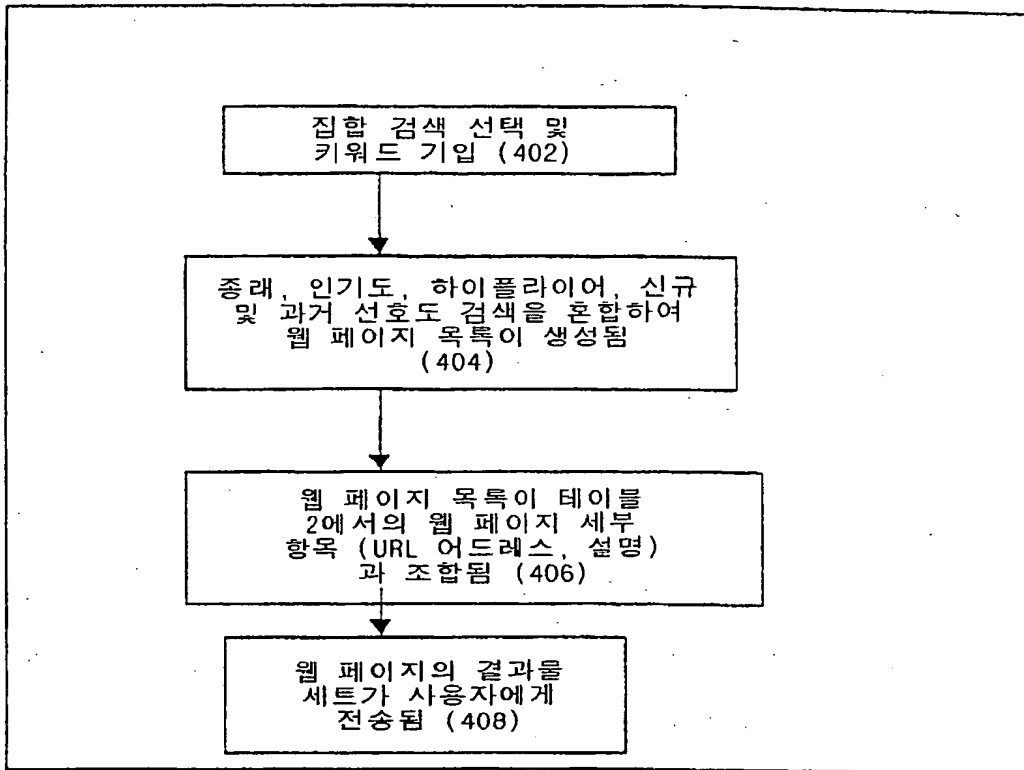
도면9



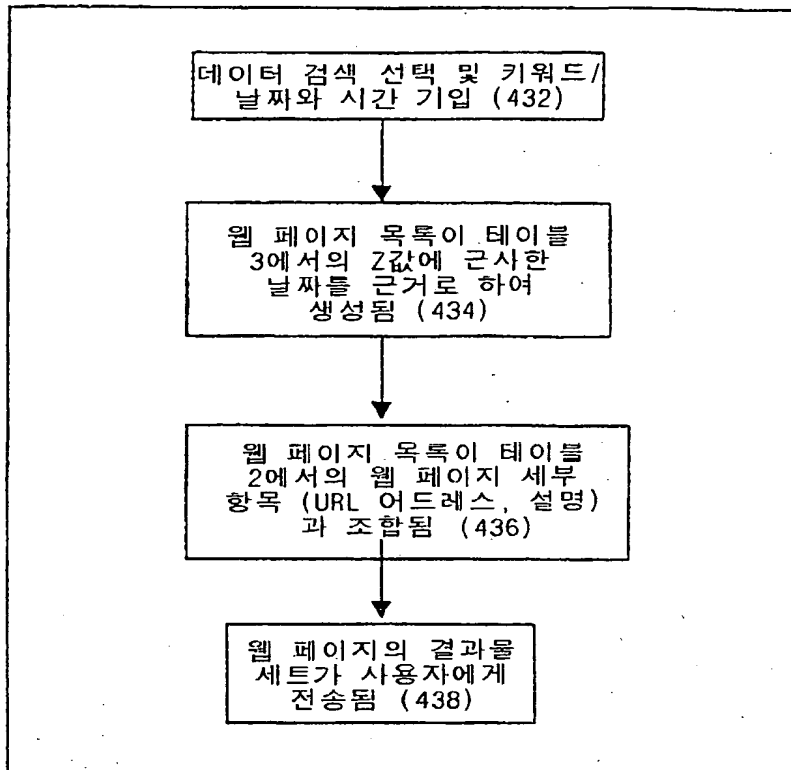
도면10



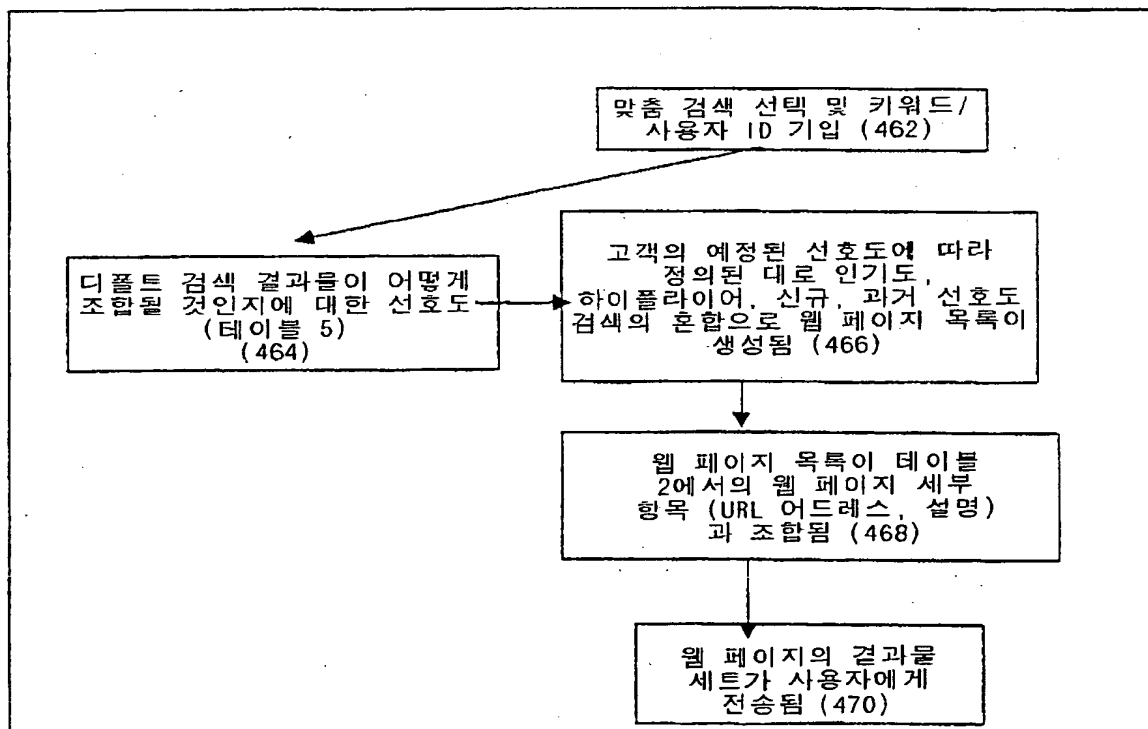
도면11



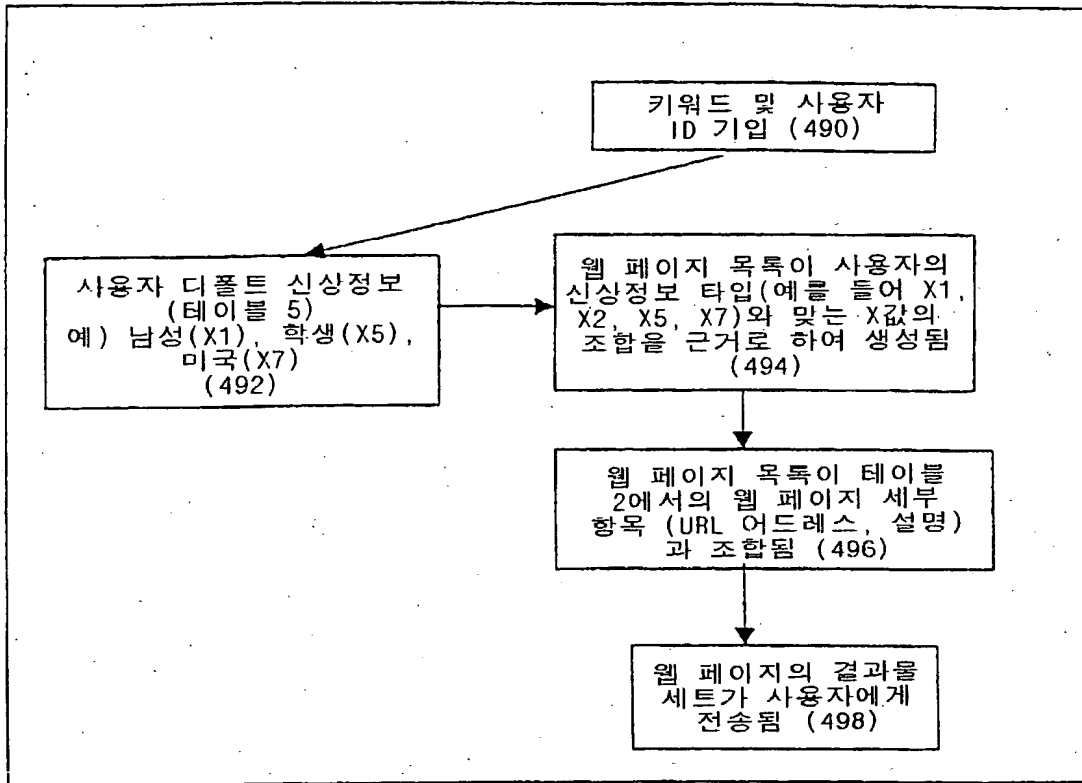
도면12



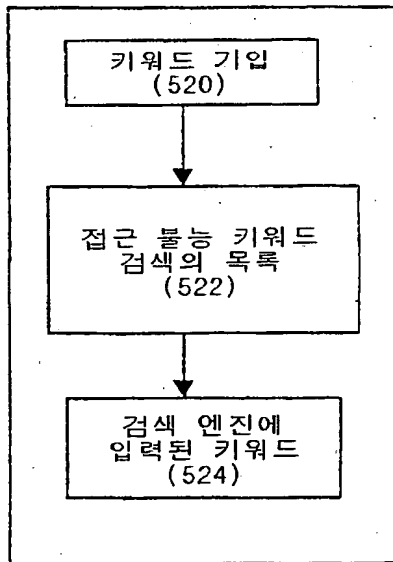
도면13



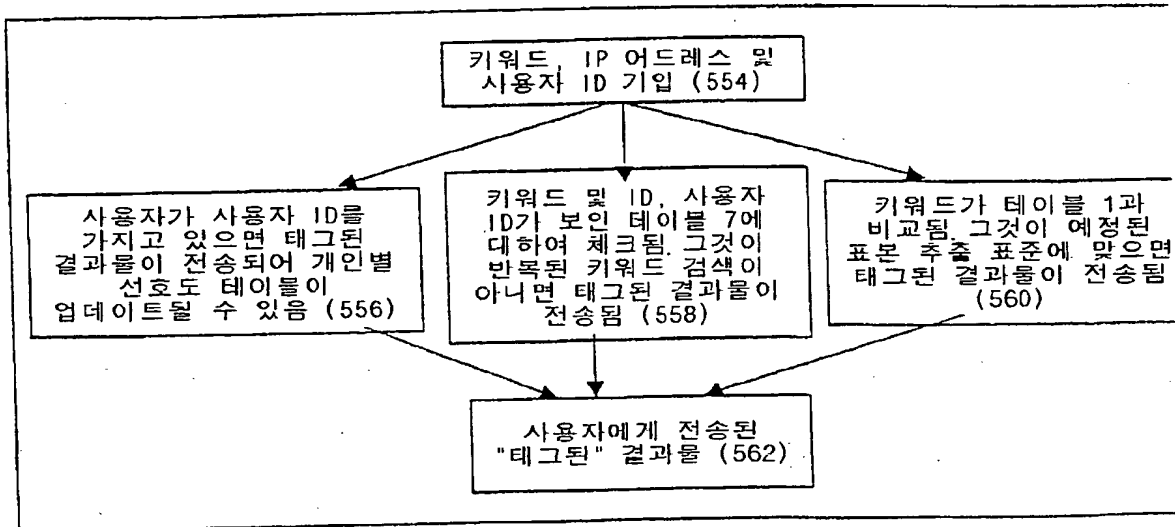
도면14



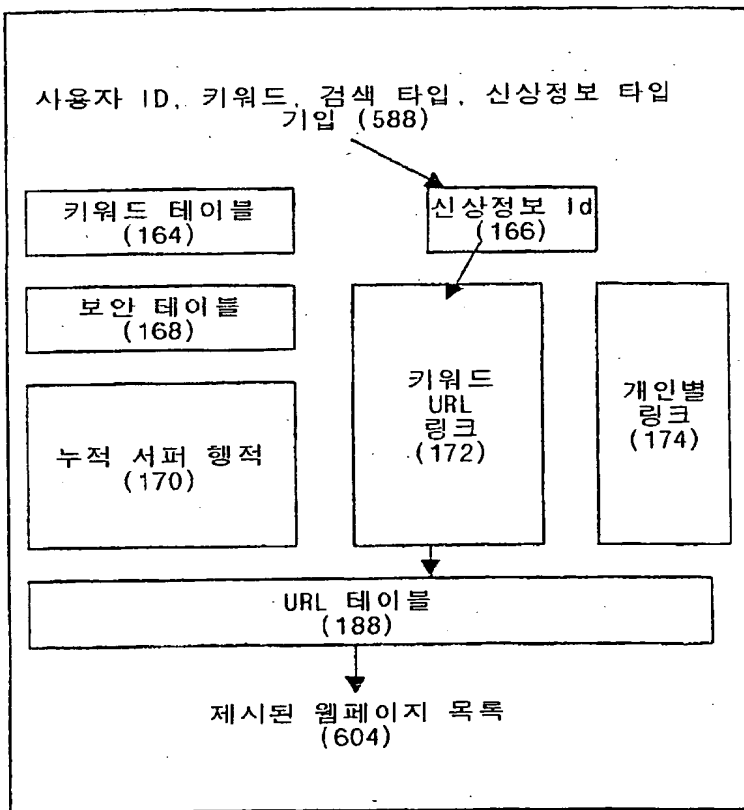
도면15



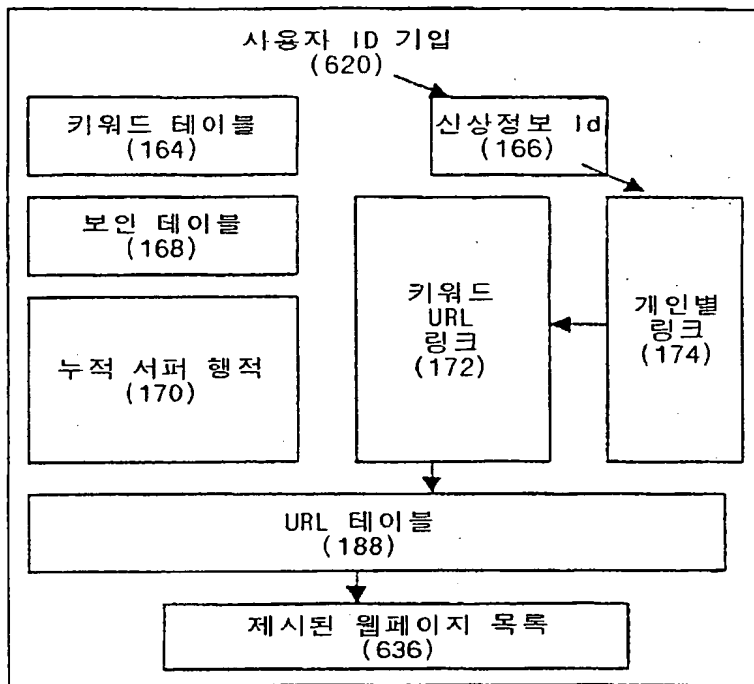
도면16



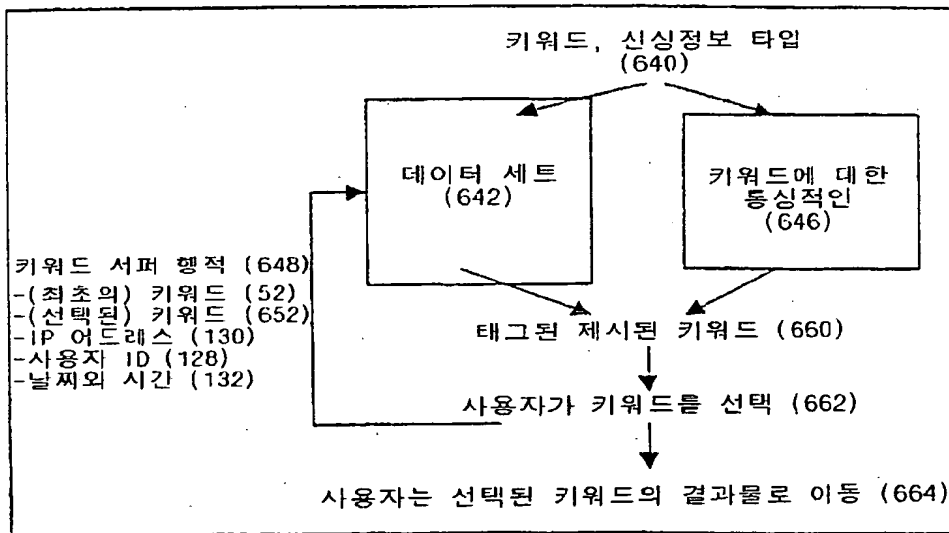
도면17



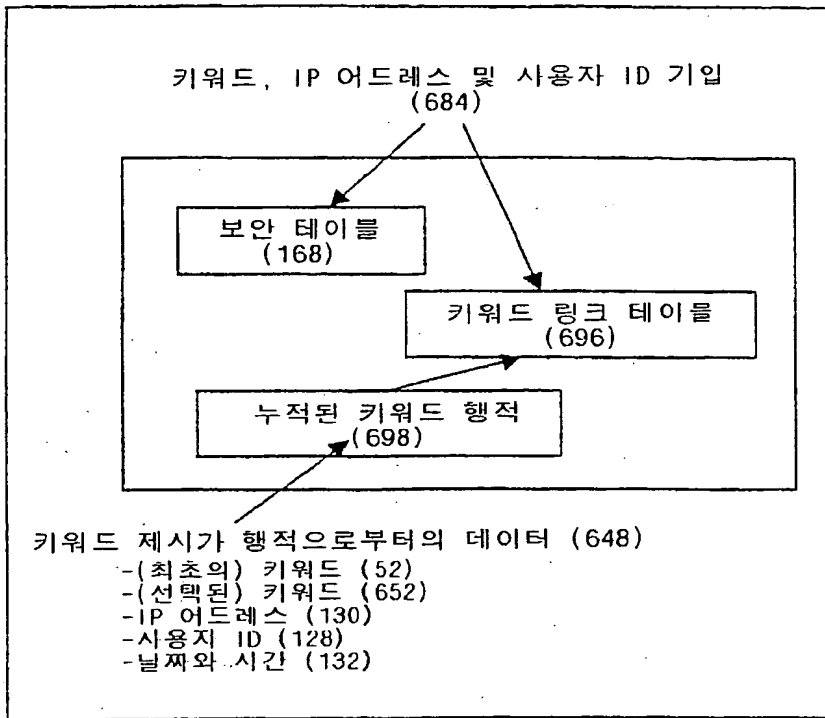
도면18



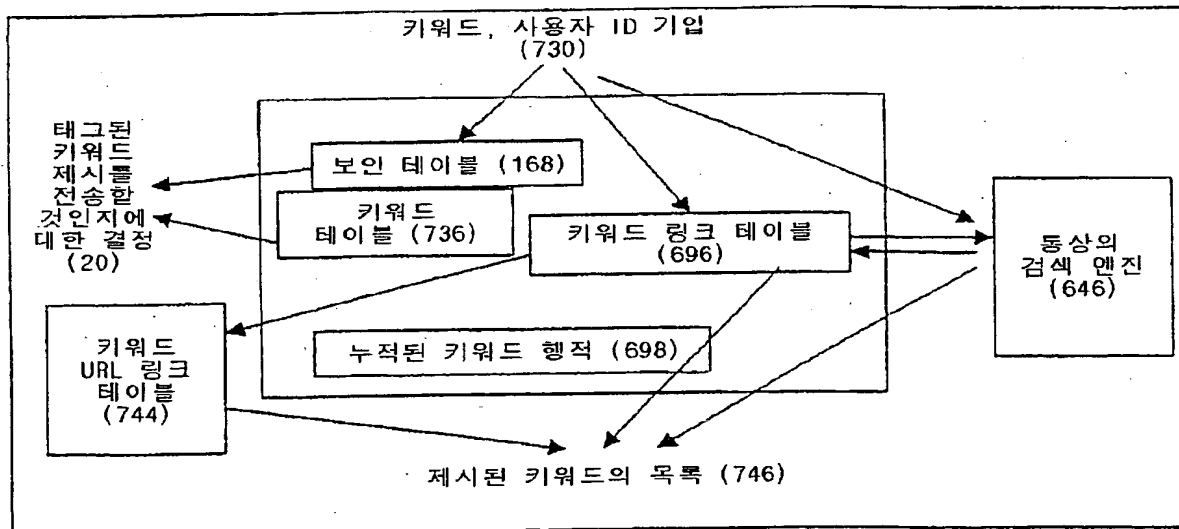
도면19



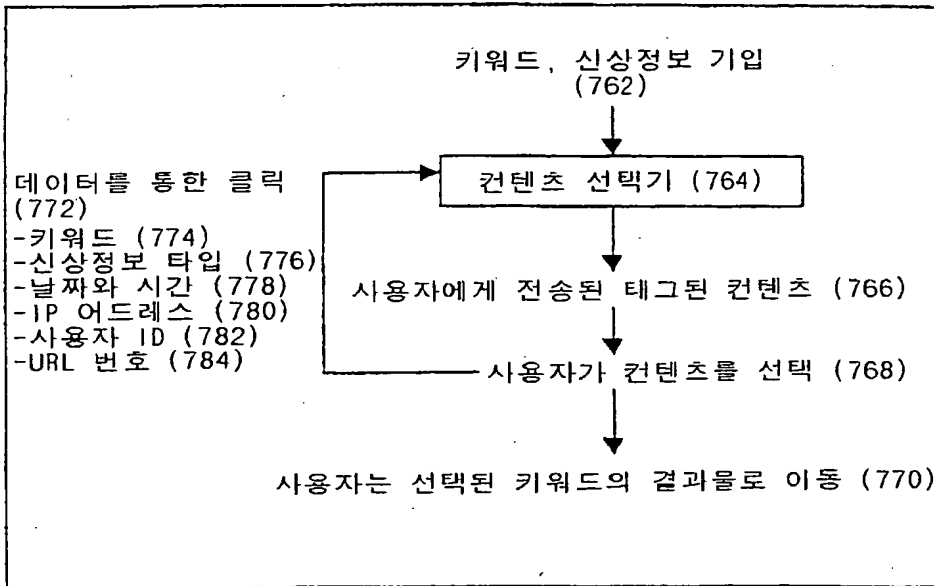
도면20



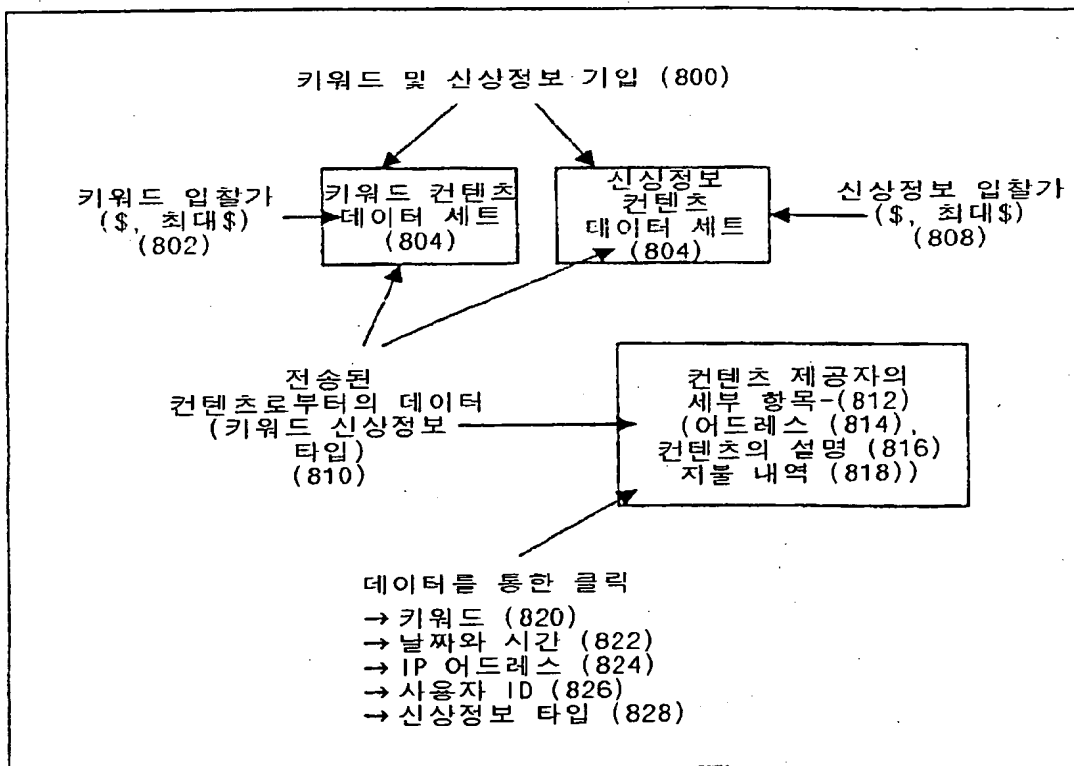
도면21



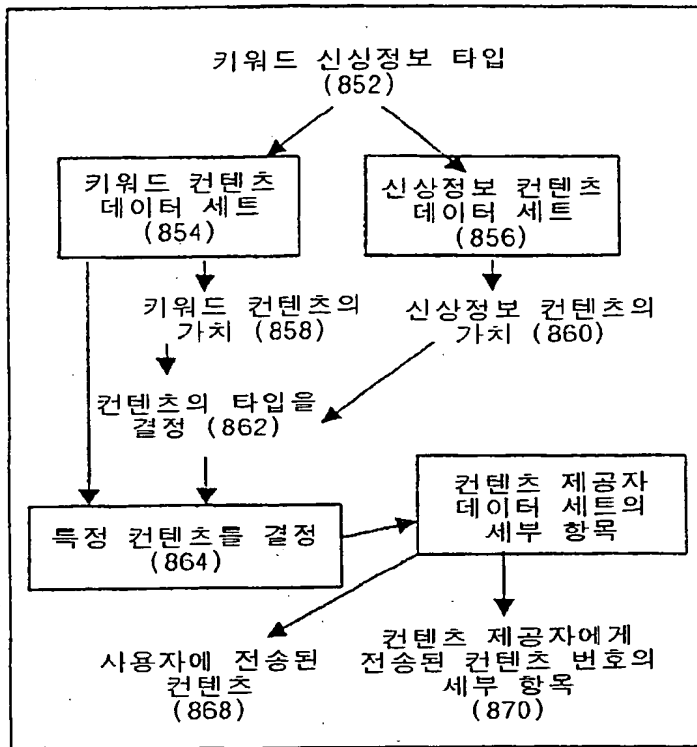
도면22



도면23



도면24



도면25

컨텐츠 (902)
웹 페이지 목록 (900)
웹 페이지 목록 (900)
웹 페이지 목록 (900)
웹 페이지 목록 (900)